PB83-165191

Evaluation of the Empirical Kinetic Modeling Approach (EKMA)

SRI International
Menlo Park, CA

Prepared for

Environmental Sciences Research Lab.
Research Triangle Park, NC

Feb 83

U.S. Department of Commerce
National Technical Information Service
NTIS

EPA-600/3-83-003
February 1983

PB83-165191

EVALUATION OF THE EMPIRICAL KINETIC MODELING APPROACH (EKMA)

by

J. R. Martinez
Atmospheric Science Center

C. Maxwell
H. S. Javitz
R. Bawol
Statistical Analysis Department
SRI International
Menlo Park, California 94025

EPA Contract 68-02-2984

Project Officer

Basil Dimitriades
Atmospheric Chemistry and Physics Division
Environmental Sciences Research Laboratory
Research Triangle Park, North Carolina 27711

ENVIRONMENTAL SCIENCES RESEARCH LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U. S. ENVIRONMENTAL PROTECTION AGENCY
RESEARCH TRIANGLE PARK, NORTH CAROLINA 27711

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO.	2.	3. RECIPIENT'S ACCESSION NO.
EPA-600/3-83-003		PB83 165191

4. TITLE AND SUBTITLE	5. REPORT DATE
EVALUATION OF THE EMPIRICAL KINETIC MODELING APPROACH (EKMA)	February 1983
	6. PERFORMING ORGANIZATION CODE

7. AUTHOR(S)	8. PERFORMING ORGANIZATION REPORT NO.
J. R. Martinez, C. Maxwell, H. S. Javitz, and R. Bawol	SRI Project 7938

9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. PROGRAM ELEMENT NO.
SRI International 333 Ravenswood Avenue Menlo Park, California 94025	CDWA1A/01 - 0649 (FY-83)
	11. CONTRACT/GRANT NO. 68-02-2984

12. SPONSORING AGENCY NAME AND ADDRESS	13. TYPE OF REPORT AND PERIOD COVERED
Environmental Sciences Research Laboratory-RTP, NC Office of Research and Development U. S. Environmental Protection Agency Research Triangle Park, NC 27711	14. SPONSORING AGENCY CODE EPA/600/09

15. SUPPLEMENTARY NOTES

16. ABSTRACT

The EKMA is a Lagrangian photochemical air quality simulation model that calculates ozone from its precursors: nonmethane organic compounds (NMOC) and nitrogen oxides (NO_x). This study evaluated the performance of the EKMA when it is used to estimate the maximum ozone concentration that can occur in an urban area and its environs. The evaluation was conducted using data for five U.S. cities: St. Louis, Houston, Philadelphia, Los Angeles, and Tulsa.

A novel statistical evaluation procedure was developed to measure the accuracy of the EKMA ozone estimates. The accuracy parameter is defined as the ratio of observed to estimated ozone. Associated with this ratio is an accuracy probability, which is defined as the probability that the ratio lies within a predefined percent (e.g. ±20 percent) of unity, a unit value of the ratio denoting perfect agreement between observation and prediction. The evaluation procedure uses NMOC and NO_x as inputs to calculate the accuracy probability of the EKMA ozone estimate. The full range of accuracy probabilities associated with the EKMA ozone estimates is displayed in graphical form on the NMOC-NO_x plane.

The report describes the results for the various cities, and discusses potential appli cations of the methodology to other models and to the assessment of ozone control strategies.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS *(This Report)*	21. NO. OF PAGES
RELEASE TO PUBLIC	UNCLASSIFIED	133
	20. SECURITY CLASS *(This page)*	22. PRICE
	UNCLASSIFIED	

EPA Form 2220-1 (Rev. 4-77) PREVIOUS EDITION IS OBSOLETE

NOTICE

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

ABSTRACT

The EKMA is a Lagrangian photochemical air quality simulation model that calculates ozone from its precursors: nonmethane organic compounds (NMOC) and nitrogen oxides (NO_x). This study evaluated the performance of the EKMA when it is used to estimate the maximum ozone concentration that can occur in an urban area and its environs. The evaluation was conducted using data for five U.S. cities: St. Louis, Houston, Philadelphia, Los Angeles, and Tulsa.

A novel statistical evaluation procedure was developed to measure the accuracy of the EKMA ozone estimates. The accuracy parameter is defined as the ratio of observed to estimated ozone. Associated with this ratio is an accuracy probability, which is defined as the probability that the ratio lies within a predefined percent (e.g. ±20 percent) of unity, a unit value of the ratio denoting perfect agreement between observation and prediction. The evaluation procedure uses NMOC and NO_x as inputs to calculate the accuracy probability of the EKMA ozone estimate. The full range of accuracy probabilities associated with the EKMA ozone estimates is displayed in graphical form on the NMOC-NO_x plane.

The report describes the results for the various cities, and discusses potential applications of the methodology to other models and to the assessment of ozone control strategies.

This report was submitted in partial fulfillment of Contract 68-02-2984 by SRI International under sponsorship of the U.S. Environmental Protection Agency. This report covers the period June 1979 to December 1981 and work was completed as of March 1982.

CONTENTS

Abstract iii

Figures vii

Tables x

1. Introduction 1
   - Objective 1
   - Background 1
   - Report Organization 1
2. Methodology 2
   - Overview of the Data Bases 3
   - Data Review, Analysis, and Selection of Days 3
   - Estimation of Ozone Using EKMA 3
   - Statistical Evaluation 6
3. EKMA Evaluation for the St. Louis Area 7
   - Data Review and Analysis 11
   - Definition of Evaluation Data Set 11
   - Evaluation of Standard EKMA 16
   - Evaluation of City-Specific EKMA 16
   - Discussion 31
4. EKMA Evaluation for the Houston Area 42
   - Data Review and Analysis 44
   - Definition of Evaluation Data Set 44
   - Evaluation of Standard EKMA 44
   - Evaluation of City-Specific EKMA 45
   - Discussion 59
5. EKMA Evaluation for the Philadelphia Area 63
   - Data Review and Analysis 66
   - Definition of Evaluation Data Set 66
   - Evaluation of Standard EKMA 66
   - Evaluation of City-Specific EKMA 69
   - Discussion 74
6. Evaluation Using Data for the Los Angeles Area 78
   - Data Review and Analysis 81
   - Definition of Evaluation Data Set 81
   - Evaluation of Standard EKMA 83
   - Evaluation of City-Specific EKMA 83
   - Discussion 97

101

7. Evaluation Using Data for the Tulsa Area..... 104
Data Review and Analysis..... 104
Definition of Evaluation Data Set..... 104
Evaluation of Standard EKMA..... 107
Evaluation of City-Specific EKMA..... 107
Discussion..... 111
8. Conclusions, Implications, and Recommendations..... 113
Discussion of Results..... 113
Implications for EKMA Applications..... 114
Recommendations..... 117
Concluding Remarks..... 118

Appendix--Equations for Standard-EKMA Ozone Estimates..... 119
References..... 120

FIGURES

Number		Page
1	Location of Regional Air Monitoring Systems (RAMS) Stations	12
2	Scatterplot of Maximum Daily Ozone Concentration as a Function of Daily Maximum Temperature for the 1976 St. Louis Data	14
3	Scatterplot of Daily Maximum Ozone Concentration as a Function of Daily Core-City 0600-0900 (CDT) Average Ratio of NMHC to NO_x for the 1976 St. Louis Data	15
4	Scatterplot of Observed Ozone and Standard-EKMA Ozone Estimate for St. Louis	22
5	Scatterplot of 0600-0900 (CDT) Concentrations of NMOC and NO_x for St. Louis	27
6	Scatterplot of OBS/EST as a Function of $1/NO_x$ for Standard-EKMA Ozone Estimate in St. Louis	28
7	Accuracy Probability Plot for Standard-EKMA Ozone Estimates for St. Louis	30
8	Plot of Constant-Z Curves on NMOC-NO_x Plane for Standard-EKMA Estimates for St. Louis	32
9	Accuracy Probability Curves for Standard-EKMA Evaluation for St. Louis	33
10	Scatterplot of City-Specific and Standard EKMA Ozone Estimates for St. Louis	36
11	Scatterplot of Observed Ozone and City-Specific EKMA Ozone Estimates for St. Louis	37
12	Accuracy Probability Plot for City-Specific EKMA Ozone Estimates for St. Louis	40
13	Plot of Constant-Z Lines on NMOC-NO_x Plane for City-Specific EKMA Estimates for St. Louis	41

Number		Page
14	Scatterplot of Observed Ozone and Standard-EKMA Ozone Estimate for HAOS Data Set.....	49
15	Scatterplot of Observed Ozone and Standard-EKMA Ozone Estimate for HOMS Data Set.....	50
16	Scatterplot of NO_x and NMOC for the HAOS Data Set.....	54
17	Scatterplot of NO_x and NMOC for the HOMS Data Set.....	56
18	Accuracy Probability for Standard-EKMA Ozone Estimates for the HAOS Data Set.....	57
19	Constant-Z Plot for Standard-EKMA Ozone Estimates for the HAOS Data Set.....	58
20	Scatterplot of Observed Ozone and City-Specific EKMA Ozone Estimates for the HAOS Data.....	61
21	Scatterplot of Observed Ozone and City-Specific EKMA Ozone Estimates for the HOMS Data.....	62
22	Accuracy Probability for City-Specific EKMA Ozone Estimates for the HAOS Data Set.....	64
23	Constant-Z Plot for City-Specific EKMA Ozone Estimates for the HAOS Data Set.....	65
24	Air Quality Monitoring Network for the Philadelphia Area.....	67
25	Scatterplot of Observed Ozone and Standard-EKMA Ozone Estimate for Philadelphia.....	70
26	Scatterplot of 0600-0900 (EDT) Concentrations of NMOC and NO_x for Philadelphia.....	72
27	Accuracy Probability Plot for Standard-EKMA Ozone Estimates for Philadelphia.....	73
28	Constant-Z Plot for Standard-EKMA Ozone Estimates for Philadelphia.....	75
29	Scatterplot of Observed Ozone and City-Specific EKMA Ozone Estimate for Philadelphia.....	77
30	Accuracy Probability Plot for City-Specific Ozone Estimates for Philadelphia.....	79

Number		Page
31	Constant-Z Plot for City-Specific EKMA Ozone Estimates for Philadelphia	80
32	Air-Quality Monitoring Network for the Los Angeles Area	82
33	Scatterplot of Observed Ozone and Standard-EKMA Ozone Estimate for Los Angeles	91
34	Scatterplot of 0600-0900 NMOC and NO_x for Los Angeles	93
35	Accuracy Probability Plot for Standard-EKMA Ozone Estimates for Los Angeles	95
36	Plot of Constant-Z Curves on NMOC-NO_x Plane for Standard-EKMA Ozone Estimates for Los Angeles	96
37	Scatterplot of City-Specific EKMA Ozone Estimates and Standard-EKMA Ozone Estimates for Los Angeles	99
38	Scatterplot of Observed Ozone and City-Specific Ozone Estimate for Los Angeles	100
39	Accuracy Probability Plot for City-Specific EKMA Ozone Estimates for Los Angeles	102
40	Plot of Constant-Z Curves on NMOC-NO_x Plane for City-Specific EKMA Ozone Estimates for Los Angeles	103
41	Map of Tulsa, Oklahoma, and Vicinity	105
42	Scatterplot of Observed Ozone and Standard-EKMA Ozone Estimate for Tulsa	108
43	Scatterplot of 0600-0900 NMOC and NO_x for Tulsa	109
44	Scatterplot of Observed Ozone and City-Specific EKMA Ozone Estimate for Tulsa	112
45	Flowchart for the Application of the EKMA as a Screening Tool for Analyzing the Impact of a Control Strategy	116

TABLES

Number		Page
1	Data Bases Used in EKMA Evaluation.....	4
2	0600-0900 (CDT) Ratios of NMHC/NO_x Concentrations for St. Louis Area.....	13
3	Evaluation Data Set for St. Louis.....	17
4	Selected 0600-0900 NMOC and NO_x Concentrations at Source-Region Monitoring Sites in St. Louis.....	24
5	Means, Standard Deviations, and Correlations for Selected Variables in the St. Louis Evaluation Data Set.....	25
6	Summary of Input Parameters for Obtaining City-Specific EKMA Ozone Estimates for St. Louis.....	34
7	Means, Standard Deviations, and Correlations for City-Specific EKMA Evaluation for St. Louis.....	39
8	Evaluation Data Set for Houston: HAOS Data.....	46
9	Evaluation Data Set for Houston: HOMS Data.....	48
10	Statistical Parameters for Selected Variables in the HAOS Data Set.....	52
11	Statistical Parameters for Selected Variables in the HOMS Data Set.....	53
12	Summary of Input Parameters for Obtaining City-Specific EKMA Ozone Estimates for Houston.....	60
13	Evaluation Data Set for Philadelphia.....	68
14	Means, Standard Deviations, and Correlations for Selected Variables in the Philadelphia Data Set.....	71
15	Summary of Input Parameters for Obtaining City-Specific EKMA Ozone Estimates for Philadelphia.....	76

Number		Page
16	Evaluation Data Set for Los Angeles.....	84
17	Means, Standard Deviations, and Correlations for Selected Variables in the Los Angeles Data Set.....	92
18	Summary of Input Parameters for Obtaining City-Specific EKMA Ozone Estimates for Los Angeles.....	98
19	Evaluation Data Set for Tulsa.....	106
20	Summary of Input Parameters for Obtaining City-Specific EKMA Ozone Estimates for Tulsa.....	110

SECTION 1

INTRODUCTION

OBJECTIVE

The objective of this study is to assess the performance of the Empirical Kinetic Modeling Approach (EKMA) when it is used to estimate the maximum ozone (O_3) concentration that could occur in an urban area and its environs. Specifically, the study quantitatively measures EKMA's ability to predict maximum O_3, defines conditions under which O_3 estimates can achieve specific accuracy levels, and examines the application of EKMA as an estimator of maximum O_3.

BACKGROUND

A common problem in air pollution control requires estimating the reduction in the level of O_3 precursors, i.e. nonmethane organic compounds (NMOC)* and nitrogen oxides (NO_x), needed to achieve a prescribed decrease in O_3 concentration. To tackle this problem, one must have a model that relates O_3 to NMOC and NO_x; the EKMA, which was developed by the U.S. Environmental Protection Agency (EPA), is one such model. The EKMA has been extensively documented, and the reader should consult the references for technical details (EPA, 1977, 1978; Dodge, 1977; Trijonis and Hunsaker, 1978; EPA, 1980). EPA has prepared and disseminated a computer program that performs the calculations that relate O_3 to NMOC and NO_x (Whitten and Hogo, 1978). The output of the computer program is a graph that contains constant-O_3 contours (O_3 isopleths) as a function of NMOC and NO_x; hereafter we will refer to this graph as the "O_3 isopleth diagram."

Basically, the EKMA takes two forms: standard and city-specific. The standard EKMA is based on conditions that prevail in the Los Angeles area; the city-specific version, as the name implies, tailors the model to a particular city. We will evaluate the performance of both forms of EKMA using data for five cities.

In a typical application, the EKMA has been used to calculate the percentage reduction in NMOC and NO_x that would be required to reduce O_3 from a high observed value to some desired lower level. Under current EPA guidelines

---

*Much of the EKMA literature refers to the organic compounds as nonmethane hydrocarbons (NMHC). This term has recently been replaced by NMOC (EPA, 1980), and we will conform to this usage throughout this report.

(EPA, 1977; 1980), such an application uses the observed ozone level and the median 0600–0900 (LDT) NMOC/NO_x ratio to find the NMOC and NO_x concentrations that the model requires to produce the observed ozone. These NMOC and NO_x concentrations then become the base data used in the calculation of percent reduction of NMOC and NO_x. Thus, in this procedure the NMOC and NO_x concentrations are not necessarily the precursor levels that prevailed when the ozone was observed. Rather, they are the presumed precursor levels defined by the model.

Our assessment of the EKMA as a predictor of maximum O_3 shifts the focus of EKMA from using observed O_3 and the NMOC/NO_x ratio to define precursor levels, as was mentioned above, to using precursor levels to estimate O_3 concentration. Thus, the EKMA could be used to estimate the O_3 level produced by a given emission control strategy that yields particular levels of NMOC and NO_x. In such applications, it would be helpful for the analyst to know whether the EKMA estimate is an upper bound for the O_3 that would actually occur, and if so, to know the accuracy of the upper bound. Our study is concerned with such questions.

REPORT ORGANIZATION

Section Two describes the evaluation methodology. The results of the EKMA evaluation using data for five cities are presented in Sections Three through Seven. Section Eight contains conclusions and recommendations. An appendix defines several equations used in the study.

SECTION 2

METHODOLOGY

OVERVIEW OF THE DATA BASES

The first step in the evaluation process identified and obtained the data bases for the five cities of interest: St. Louis, Missouri; Houston, Texas; Philadelphia, Pennsylvania; Los Angeles, California; and Tulsa, Oklahoma. For each city, the model assessment procedure consisted of three steps:

- Review and analyze data and select days for the evaluation data set.
- Obtain ozone estimates using both standard and city-specific EKMA.
- Conduct a statistical evaluation of EKMA performance as a predictor of maximum O_3.

Table 1 lists the five cities and briefly describes each data base. Both air quality and meteorological data were included in each data base, but the quantity of data and the spatial coverage are by no means uniform for all the data bases. Because the data for four of the cities were obtained in special studies, the spatial coverage of pollutant patterns for these cities is more thorough than usual. For Houston in 1978, and for Tulsa and Philadelphia, the duration of the monitoring program was short, and this, in combination with other data selection criteria, resulted in small data sets being used in the EKMA evaluation.

Routinely collected monitoring data were used for Los Angeles, but this is a rich data base because the monitoring network is extensive. Los Angeles data for the period 1976-1978 were examined; 1978 was selected because it contained the highest ozone levels.

The sections below describe the steps in the evaluation process.

DATA REVIEW, ANALYSIS, AND SELECTION OF DAYS

The objectives of the data review and analysis process are to identify the available data and to define the criteria for selecting the days for the EKMA evaluation data set.

Each data base was examined to determine what correlations, if any, exist between high ozone concentrations and other pollutants and meteorological parameters. These correlations were used to define criteria for selecting days to test the city-specific and the standard EKMA. We seek to identify

TABLE 1. DATA BASES USED IN EKMA EVALUATION

City	Date	Data Base			Reference
		Type	Title	Source	
St. Louis, Missouri	May-Oct 1976	Special	Regional Air Pollution Study (RAPS)	EPA	Cardwell (1980)
Houston, Texas	May-Oct 1977	Special	Houston Area Oxidant Study (HAOS)	SRI	Ludwig and Martinez (1979) Ludwig et al. (1979)
	15 Sep - 12 Oct 1978	Special	Houston Oxidant Modeling Study (HOMS)	EPA	Martinez (1982) Nitz and Martinez (1982)
Philadelphia, Pennsylvania	Jul-Aug 1979	Special	Philadelphia Oxidant Data Enhancement Study	EPA	Westberg and Sweeny (1980)
Los Angeles, California	May-Oct 1978	Routine Monitoring		California Air Resources Board	--
Tulsa, Oklahoma	Jul-Sep 1977	Special		EPA	Eaton and Dimmock (1979) Eaton et al. (1979)

precursor and meteorological conditions associated with ozone concentrations that equal or exceed 100 ppb; these conditions are necessary, but may not be sufficient, for such concentrations of ozone to occur. (That is, high ozone concentrations never occur in the absence of such conditions, but may not occur even if the conditions are present.) This approach yields a set of selected modeling days that encompasses the high-ozone-concentration events, but that also contains a number of low-ozone cases. Although the latter may strain the worst-case assumptions of the EKMA, we have retained them to extend the range of ozone concentrations in the evaluation.

The data were also analyzed to identify monitoring sites within the urban area that can be considered to be good indicators of 0600-0900 (LDT) NMOC and NO_x concentrations, which EKMA considers to be the principal cause of the ozone maximum that occurs downwind of the source region. We will refer to these sites as source-region monitors.

The initial step of the analysis was to compile daily summaries for each data set. The daily summary provides information on pollutant and meteorological data, including the date, the maximum ozone concentration observed throughout the monitoring area, the station where the peak ozone concentration was observed, and the first or only hour when the peak was recorded. The daily summaries also include average 0600-0900 NMOC and NO_x concentration for the source-region monitors and the ratio of the two averages. Additional information is listed indicating whether the selected source-region monitors had missing data for the times for which the average NMOC and NO_x concentrations were calculated. Meteorological data contained in the daily summaries include the daily maximum temperature, the morning minimum temperature, the 0600-1400 (LDT) average wind speed and direction, and the standard deviations of the wind speed and direction at one selected rural location, if available. Cloud cover data for 1200 (LDT) and any available solar radiation data are also listed. Estimates of the background ozone concentrations, which would include transport effects, are listed for each day, but these were not available for all the cities.

Several restrictions were imposed on the data used for calculating the average 0600-0900 NMOC and NO_x concentrations at the source-region monitors. At each individual monitoring site, an average 0600-0900 NMOC and NO_x concentration would be calculated only if no less than two of the possible three hours of pollutant data were available; otherwise, the 0600-0900 concentration was considered to be missing. The composite (spatial) average of the 0600-0900 pollutant concentrations would be calculated if averages existed for at least half of the source-region monitors; otherwise the spatial average was considered to be missing.

The daily average wind speeds and directions were calculated for the hours from 0600-1400 (LDT). These hours were selected to indicate the general transport between the hours of peak NMOC and NO_x emissions and the time of the peak ozone concentrations.

The background (including transported) ozone concentration was determined as follows. Based on the average 0600-1400 wind direction, the background ozone concentration was assumed to equal the peak* ozone concentration recorded at the upwind station most distant from the urban area. This method of determining background ozone differs from the method suggested in the EKMA literature, which uses the 1100-1300 average ozone at an upwind location (EPA, 1977). It is recognized that as the transported ozone enters the urban area its level can be reduced due to reaction with nitric oxide; however, the amount of ozone scavenging is not well defined. By using the peak upwind ozone, we are conservatively assured of having the worst-case background ozone level.

The essential criteria for selecting days for the evaluation data set comprised:

- Data availability
- The time of occurrence of the daily ozone maximum
- The prevalence of meteorological conditions that are necessary, but may not be sufficient, for the O_3 level to be at least 100 ppb.

The 100-ppb cutoff was chosen to ensure that all days when the ozone national ambient air quality standard (NAAQS) of 120 ppb was exceeded are included in the evaluation. Such days should be included because EKMA is based on worst-case assumptions and cannot be expected to perform well for days that do not meet these conditions. Because we are interested in maximum ozone potential with reference to an air quality standard, days with conditions associated with low ozone concentrations are not critical to our evaluation.

ESTIMATION OF OZONE USING EKMA

Estimates of O_3 concentration were obtained using both standard and city-specific EKMA. The standard-EKMA estimates were computed using equations fitted by Holton (1980) to the standard O_3 isopleth diagram. These equations were independently tested at SRI. The test verified that the fit was very good, but a slight bias was detected and corrected by means of supplemental equations. Holton's formulas and the adjustment equations are presented in the appendix of this report.

The city-specific ozone estimates were obtained using the EKMA computer code. For each city, the NMOC and NO_x precursors in the evaluation data set were the model's inputs, and the code computed the ozone maximum using the CALCULATE option (Whitten and Hogo, 1978).

---

*As used in this report, the term "peak ozone" refers to the maximum 1-hour average ozone.

For both standard and city-specific EKMA, the inputs used to estimate O_3 were the 0600-0900 (LDT) NMOC and NO_x, spatially averaged over all the source-region monitors. Of course, the same NMOC and NO_x coordinates were used to obtain both standard and city-specific O_3 estimates.

STATISTICAL EVALUATION

Approach

The evaluation is designed to test whether EKMA can be used to estimate an upper bound for the maximum ozone concentration that can occur in a given geographical region. To test this usage of EKMA we must define the daily maximum ozone concentration.

Daily Maximum--The maximum 1-hr average O_3 concentration that occurs at some time and location during a day in an urban region: The observed maximum value recorded at a monitoring site, which may differ from the true daily maximum. (The latter may occur in a location where there is no monitor.) This is the variable with which the EKMA ozone estimate will be compared.

The question of EKMA performance is, then, How well do EKMA estimates serve as an upper bound for observed daily maximum ozone concentrations? EKMA estimates are an upper bound if they are equal to or higher than the observed maximum on every day. Denoting the observed daily O_3 maximum by OBS and the EKMA estimate by EST, then we want to test whether the inequality $OBS \leq EST$ holds. However, this criterion alone is inadequate, because it does not indicate whether an upper bound is higher than necessary: An excessively high upper bound could lead to unduly conservative decisions about allowable levels of precursors. Both underprediction and overprediction must be considered in evaluating the performance of the EKMA. But, as we will show, it is appropriate to relax the strict requirement that $OBS \leq EST$ to allow a certain amount of underprediction. Such considerations are discussed below.

In evaluating the performance of the EKMA, we must recognize that the measurements of O_3, NMOC, and NO_x are not exact; they are inevitably corrupted by random and systematic errors. Recent analyses of the quality of pollutant measurements made in monitoring programs in the Houston area indicate that most of the measurement errors for O_3, NMOC and NO_x fall within an interval of ±20 percent (Ludwig et al., 1979; Martinez, 1981). Thus, even if EKMA were perfect, O_3 estimates and observations would be expected to differ. Consequently, using the ±20 percent figure, we define the following regions of varying accuracy for the ratio of observed daily maximum to estimated O_3 (denoted by OBS/EST):

- Region 1: $1.2 < OBS/EST$
- Region 2: $0.8 \leq OBS/EST \leq 1.2$
- Region 3: $OBS/EST < 0.8$.

Region 1 contains ratios representing cases of substantial underprediction. Such cases are of special concern because they represent instances

where the EKMA does not produce an upper bound for the observation. Ratios in Regions 2 and 3 satisfy the inequality OBS/EST $\leq$ 1.2, which allows for the possibility of an apparent underestimation of at most 20 percent. Such underestimation could be the result of measurement error, and we consider it acceptable. Consequently, in testing whether EKMA estimates are upper bounds for observed O_3, we will relax the requirement OBS $\leq$ EST, and substitute OBS $\leq$ 1.2 EST instead. We consider ratios in Region 2 to represent the most accurate predictions that can be made, because the error is at most 20 percent. Region 3 contains ratios representing cases of substantial overprediction. Such cases would be of concern in applications of EKMA only if the estimate is over 120 ppb, which is the NAAQS.

To evaluate EKMA, we devised a method for calculating the likelihood that a given EKMA estimate falls in one of the three regions defined above. In general, the evaluation procedure consists of three steps:

Step 1: Compare observed (OBS) and estimated (EST) ozone.

Step 2: Investigate relationships between OBS/EST ratio and NMOC, NO_x, and other variables.

Step 3: Define regions on the NMOC-NO_x plane where the EKMA estimates attain prescribed levels of predictive accuracy.

Descriptions of these three steps are presented below.

Step 1: Compare Observed and Estimated Ozone

To compare observed and estimated ozone, we began by plotting a scatter diagram of OBS as a function of EST. We could then define the three regions of varying OBS/EST ratio and examine the location of the points with respect to the three regions. In addition, we investigated the correlation, if any, between OBS and EST.

Step 2: Investigate the Relationships Between the Variables

In this step we examined the correlation between OBS/EST and NMOC, NO_x, and other variables such as temperature and background ozone. We also investigated the correlation between NMOC and NO_x, and between pairs of the other variables. When possible, a multiple regression equation was derived using stepwise regression that relates the OBS/EST ratio to a subset of the variables of interest. This multiple regression equation was used in the third step of the analysis to define regions of prescribed accuracy on the O_3 isopleth diagram.

In preparing for the third step of the evaluation, we used the correlation, if any, between NMOC and NO_x to define an evaluation region on the O_3 isopleth diagram. (The evaluation region is defined to limit the results of the EKMA evaluation to the portion of the O_3 isopleth diagram where the data are located.) The evaluation region was defined by regressing NO_x on NMOC and placing a band around the regression line so as to capture approximately 95 percent of the joint distribution of NMOC and NO_x. Thus, the evaluation

region is defined by $NO_x = a + b\ (NMOC) \pm 1.96s$, where a and b are the coefficients of the regression line, and s is the standard error of estimate for the regression. The evaluation region may be modified further as a result of criteria used to select the evaluation data set.

Step 3: Define Regions of Predictive Accuracy

Because the performance of the EKMA model could vary as a function of NO_x, NMOC, and other variables (e.g. temperature and background ozone), we sought to estimate the likelihood that the ratio OBS/EST falls in Region 1, 2 or 3 as a function of these variables. Our approach is as follows.

The ratio OBS/EST is used as the dependent variable in a multiple linear regression equation of the form

$$OBS/EST = a_o + \sum_{i=1}^{N} a_i X_i = F(X)$$

where

X_i = independent variable, such as NO_x, NMOC, temperature

a_i = regression coefficient

N = number of independent variables.

and F(X) is shorthand notation for the regression equation. The standard error of estimate, s, is also obtained for this regression.

Define R = OBS/EST, $R_1 = 1.2$, and $R_2 = 0.8$. We wish to find the probability that $R > 1.2$, or $0.8 \leq R \leq 1.2$, or $R < 0.8$. Let $P(R \leq R_j)$ = probability that $R \leq R_j$, $j = 1,2$. Then

$$P(R \leq R_j) = \Phi\left[\frac{R_j - F(X)}{s}\right], \quad j = 1,2$$

where Φ is the cumulative normal probability distribution function. Then for the three accuracy regions we have

- Region 1:

$$P(R > 1.2) = 1 - P(R \leq R_1) \qquad (1a)$$

- Region 2:

$$P(0.8 \leq R \leq 1.2) = P(R \leq R_1) - P(R \leq R_2) \qquad (1b)$$

- Region 3:

$$P(R < 0.8) = P(R \leq R_2) \quad . \qquad (1c)$$

(Note that the above equation for $P(R \leq R_j)$ is general and thus valid not only for R_j equal to 1.2 and 0.8, but also for any other values of R_j).

The variables R_i include either NO_x or NMOC or both (or a transformation thereof, e.g. $1/NO_x$) and possibly other variables such as temperature. In the most general case, let $R_1 = f(NO_x)$, $R_2 = g(NMOC)$ and R_k, $k = 3, \ldots, N$, where f and g denote transformations of NO_x and NMOC, respectively. Then we can calculate and plot the probabilities in Eq. 1 as functions of $(R_1R_1 + R_2R_2)$ for fixed values of R_k. The plot allows us to estimate the probability that for a given (NMOC, NO_x) pair the ratio OBS/EST falls in one of the three accuracy regions in Eq. 1. The accuracy regions and their probabilities are then identified on the ozone isopleth diagram.

For EKMA to provide an upper bound for the maximum ozone, we require that the ratio OBS/EST have a high probability of being in Region 2 or 3. Conversely, EKMA fails to produce an upper bound if there is a high probability that the ratio falls in Region 1. The most desirable outcome is to have a high probability that the ratio falls in Region 2. We still get an upper bound if the ratio falls in Region 3, but the bound is not sharp, because the overprediction would be greater than 20 percent.

A question remains about what values would be acceptable for the probabilities associated with the three accuracy regions. We have purposely refrained from specifying any probability thresholds at this time, and will let the data speak for themselves instead.

SECTION 3

EKMA EVALUATION FOR THE ST. LOUIS AREA

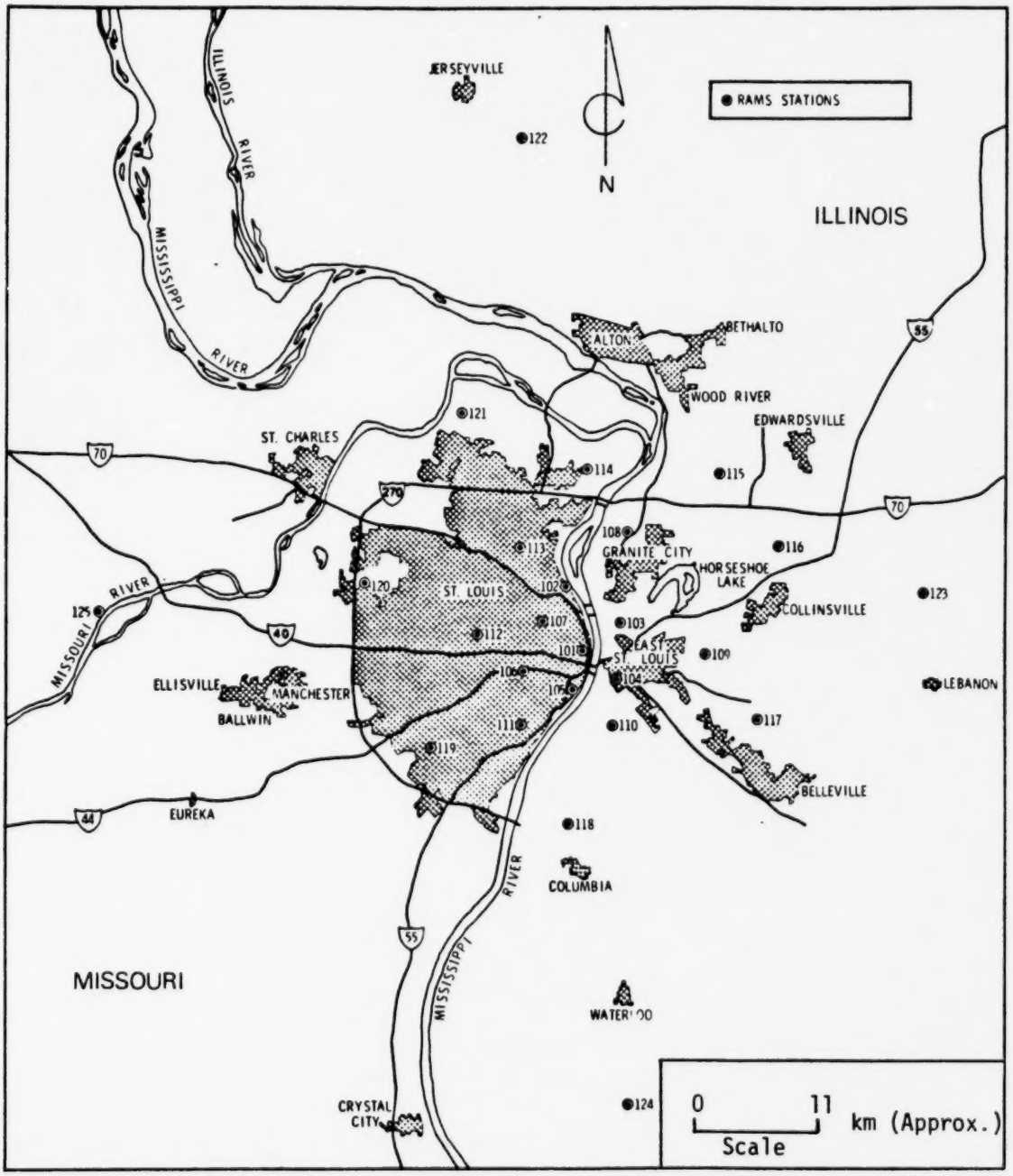
DATA REVIEW AND ANALYSIS

The St. Louis data were collected during the Regional Air Pollution Study (RAPS) from May through October 1976. Twenty-five monitoring stations were used during RAPS; their locations are shown in Figure 1. Four source-region monitors (Sites 101, 105, 106, and 107) were selected to determine spatially averaged 0600-0900 (CDT) NMOC and NO_x concentrations. Meteorological data used in the analysis were collected at a rural site south of St. Louis (Site 124).

Six monitoring stations that measured NMOC and NO_x were investigated to determine whether they should be classified as source-region sites. The six stations are 101, 102, 105, 106, 107 and 111. Figure 1 shows that all are within the metropolitan area. In relation to the downtown area, Sites 101 and 105 are closest, Stations 106 and 107 are about 6 km west, Station 102 is approximately 7 km north-northwest, and Station 111 is about 8 km south-southwest. Table 2 compares the 0600-0900 (CDT) NMOC/NO_x ratios at these sites, showing range, mean, standard deviation, and median ratio. It is evident from Table 2 that Stations 102 and 111 have similar ratios, which tend to be lower than at the other four stations. Thus, Stations 101, 105, 106, and 107, which are closest to downtown St. Louis, show the highest mean and median ratios; the difference in their mean ratios is not statistically significant. Consequently, stations 101, 105, 106, and 107 were selected as the source-region monitors that determine the spatially averaged 0600-0900 (CDT) NMOC and NO_x concentrations.

To define the necessary conditions for high-ozone-concentration days, we analyzed the relationship between ozone and precursor and meteorological variables. Figure 2 plots the maximum observed daily ozone concentration against the maximum daily temperature recorded at the rural site. Overall, there is a good linear correlation ($r = 0.7$). The data distribution in Figure 2 suggests that the envelope of the points may have an exponential form. Of particular interest is that no peak ozone concentrations exceeding 100 ppb occurred when the maximum temperature was below 24.0°C. However, as befits a necessary but not sufficient condition, Figure 2 shows that a number of ozone maxima below 100 ppb were recorded on occasions when the maximum temperature exceeded 24°C.

The distribution of the maximum ozone concentrations as a function of the ratio of the average 0600-0900 NMOC and NO_x concentrations recorded at the source monitors is shown in Figure 3. Ozone concentrations exceeding 100 ppb occurred only when the NMOC/NO_x ratio exceeded 5.0.



SOURCE: STROTHMAN AND SCHIERMEIER, 1979

Figure 1. Location of Regional Air Monitoring Systems (RAMS) stations.

TABLE 2. 0600-0900 (CDT) RATIOS OF NMHC/NO_x CONCENTRATIONS FOR ST. LOUIS AREA

Station Number	Range of Ratios	Mean	Standard Deviation	Median	Number of Ratios
101	0.5-61.6	9.4	7.4	8.1	110
102	0.2-39.0	5.5	4.1	5.3	104
105	2.7-26.1	8.5	4.1	7.3	89
106	0.3-56.4	9.2	6.0	7.9	121
107	2.2-29.1	9.2	3.6	8.5	149
111	0-79.2	5.1	7.3	4.4	118

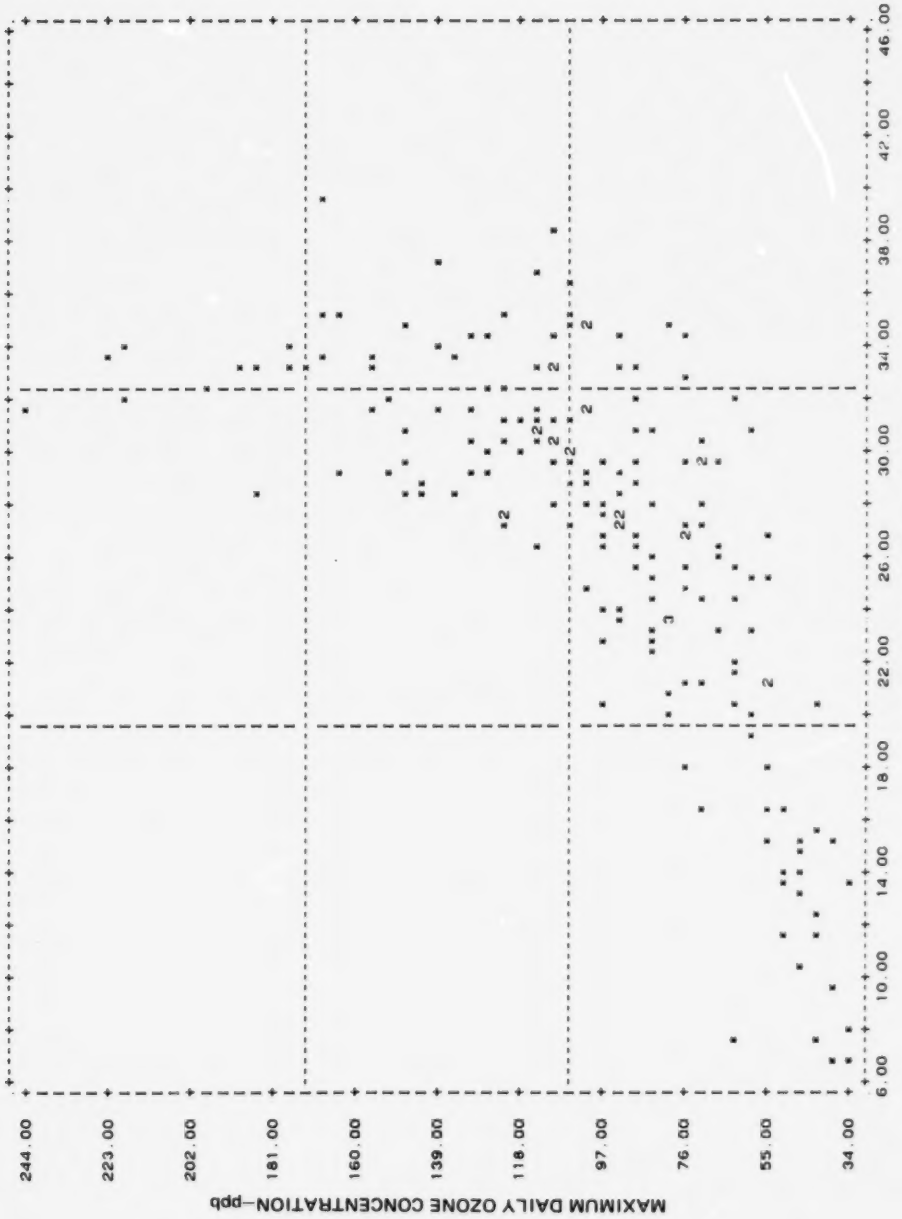


Figure 2. Scatterplot of maximum daily ozone concentration as a function of daily maximum temperature for the 1976 St. Louis data.

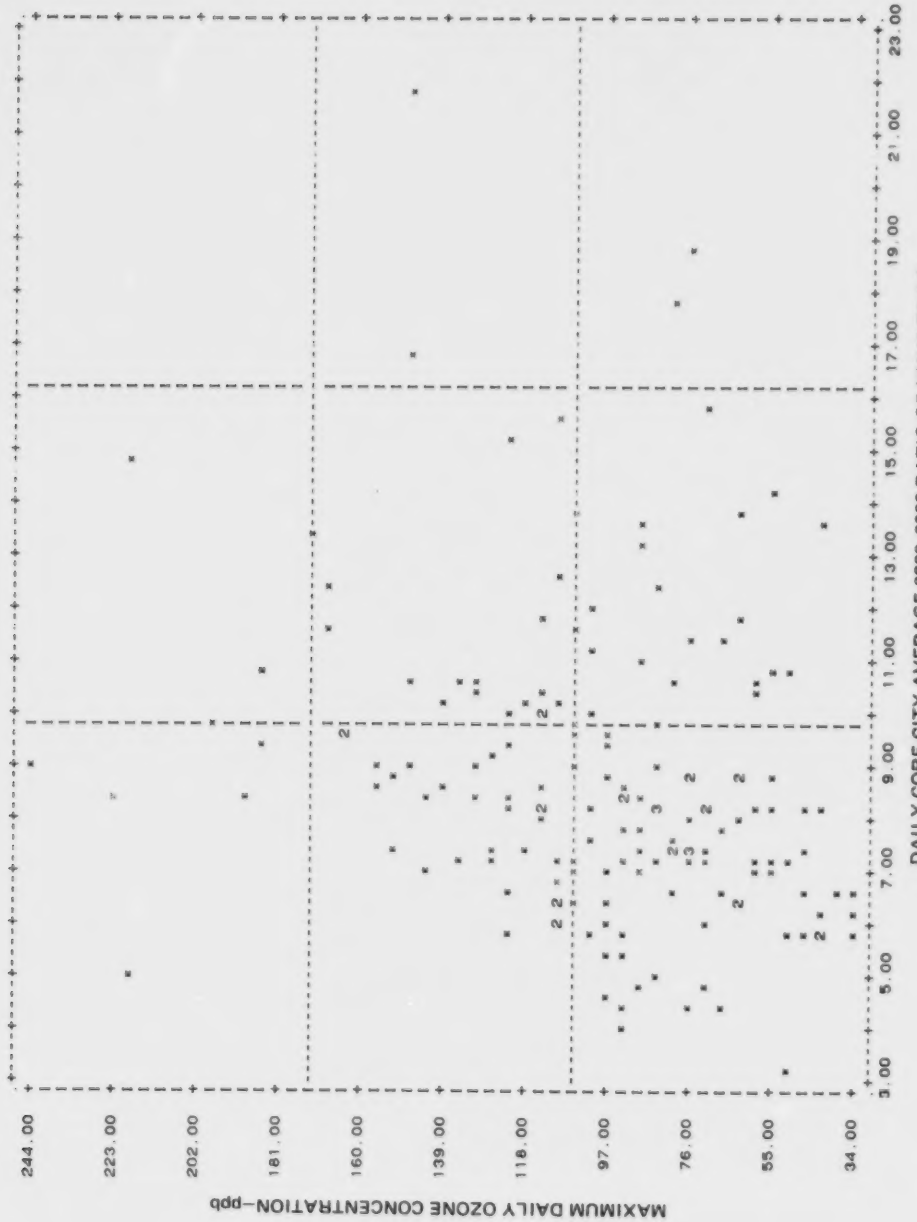


Figure 3. Scatterplot of daily maximum ozone concentration as a function of daily core-city 0600–0900 (CDT) average ratio of NMHC to NO_x for the 1976 St. Louis data.

We also investigated several other variables to see whether they imposed any restrictions on the occurrence of ozone levels above 100 ppb. Variables examined included cloud cover, dew point temperature, and wind speed and direction. For cloud cover, it was found that although the highest ozone levels occurred on days when cloud cover was low, ozone concentrations exceeding 120 ppb were observed for all cloud cover conditions. Accordingly, cloud cover was not used as a selection criterion. Similar negative findings apply to dew point and wind direction.

DEFINITION OF EVALUATION DATA SET

Based on the analysis discussed previously, the selection criteria are:

- Maximum temperature higher than 24°C
- Ratio of NMOC to NO_x greater than 5.0.

Two additional criteria were applied:*

- Time of maximum ozone $\geq$ 1200 (LT)--A daily maximum that occurs before noon indicates an unusual event that probably cannot be predicted by EKMA (EPA, 1977).
- Concurrent daily maximum O_3 and 0600-0900 NMOC and NO_x spatially averaged over the source-region monitors are required. At a given monitor, 0600-0900 NMOC or NO_x was considered missing if data for two hours were missing. The spatially averaged 0600-0900 NMOC or NO_x was considered to be missing if more than half of the monitors have missing data.

The evaluation data set for St. Louis is listed in Table 3, which includes the standard and city-specific ozone estimates. The number of cases in the set is 100.

EVALUATION OF STANDARD EKMA

Comparison of Observed and Estimated Ozone

Figure 4 shows a scatterplot of OBS as a function of EST. Of the 100 points plotted, four are in Region 1, 18 in Region 2, and 78 in Region 3. Thus, 4 percent of the cases are underpredicted and the remainder satisfy the inequality OBS $\leq$ 1.2 EST.

The observed and estimated O_3 depicted in Figure 4 are correlated: the correlation coefficient is $r = 0.49$, and is highly statistically significant ($p < 0.0001$). However, there is a fair amount of scatter. The regression line relating OBS and EST, shown in Figure 4, is defined by $\hat{Y} = 0.232$ EST + 74.4, where the standard error of estimate is $s = 32.4$, all the units are ppb,

*These two criteria also apply to the other four cities.

TABLE 3. EVALUATION DATA SET FOR ST. LOUIS

Date (1976)	Precursors NMOC (ppbC)	Precursors NO_x (ppb)	Maximum Temperature (°C)	Background O_3 (ppb)	Observed Maximum Ozone Station	Observed Maximum Ozone Time (CDT)	Observed Maximum Ozone Concentration (ppb)	EKMA Ozone Estimate (ppb) Standard	EKMA Ozone Estimate (ppb) City-Specific
11 May	265	41	27.5	74	24	1700	100	117	108
15 May	324	36	24.9	80	24	1500	80	124	112
22 May	318	37	30.7	98	25	1500	118	124	112
24 May	383	65	24.1	76	24	1700	96	151	126
29 May	882	83	28.5	106	21	1400	139	230	174
30 May	171	24	29.6	76	14	1600	100	86	92
31 May	403	48	26.3	47	23	1300	89	146	125
1 Jun	425	47	29.6	98	24	1400	151	148	126
2 Jun	435	50	29.1	82	18	1700	93	152	128
3 Jun	508	65	27.7	86	24	1300	94	174	140
4 Jun	471	48	28.1	98	22	1500	98	155	130
5 Jun	465	34	28.8	60	22	1700	91	136	120
6 Jun	504	50	30.9	107	14	1800	118	160	134
7 Jun	1856	187	32.5	137	22	1600	198	388	261
8 Jun	1111	75	33.9	115	15	1700	221	237	176
9 Jun	615	59	31.2	88	23	1200	121	181	146
11 Jun	601	66	34.4	91	15	1500	133	186	148
14 Jun	807	95	33.4	84	14	1800	93	232	174
19 Jun	186	28	26.2	55	24	1500	69	92	95
21 Jun	1436	175	25.3	67	19	1400	86	345	240
22 Jun	947	145	28.1	86	21	1500	112	274	197
25 Jun	806	108	32.2	65	11	1200	154	239	179

TABLE 3 (continued)

Date (1976)	Precursors NMOC (ppbC)	Precursors NO_x (ppb)	Maximum Temperature (°C)	Background O_3 (ppb)	Observed Maximum Ozone Station	Observed Maximum Ozone Time (CDT)	Observed Maximum Ozone Concentration (ppb)	EKMA Ozone Estimate (ppb) Standard	EKMA Ozone Estimate (ppb) City-Specific
26 Jun	552	76	33.6	107	22	1600	138	187	147
27 Jun	354	44	37.0	80	15	1300	116	135	118
28 Jun	698	63	33.3	58	16	1200	90	193	153
29 Jun	533	66	29.7	57	22	1700	74	178	143
30 Jun	299	36	25.2	54	24	1600	60	120	110
1 July	697	103	30.6	61	9	1600	112	221	167
4 July	225	28	29.7	67	20	1200	89	100	99
7 July	504	66	31.5	80	14	1500	103	174	140
8 July	289	34	34.0	73	18	1700	141	117	108
9 July	357	58	33.2	73	14	1200	112	143	122
10 July	275	23	24.8	50	15	1400	102	101	101
12 July	146	13	34.8	60	10	1600	105	68	83
13 July	228	45	32.2	106	14	1600	223	121	103
17 July	691	85	27.9	55	5	1600	85	211	162
18 July	587	70	30.6	102	21	1600	126	188	149
19 July	951	56	30.9	83	22	1300	148	202	156
20 July	448	60	34.3	63	14	1200	78	161	133
21 July	337	46	34.7	51	15	1300	108	134	117
22 July	487	68	36.6	73	16	1600	107	172	139
23 July	451	44	37.4	56	23	1500	141	148	127
24 July	147	14	38.4	87	18	1600	112	70	84
25 July	324	48	35.1	93	20	1500	126	133	116
27 July	516	69	30.1	69	22	1700	121	177	142
29 July	383	44	29.1	145	20	1600	155	140	121
30 July	1216	98	35.1	93	8	1300	170	270	196

TABLE 3 (continued)

Date (1976)	Precursors NMOC (ppbC)	Precursors NO_x (ppb)	Maximum Temperature (°C)	Background O_3 (ppb)	Observed Maximum Ozone Station	Observed Maximum Ozone Time (CDT)	Observed Maximum Ozone Concentration (ppb)	EKMA Ozone Estimate (ppb) Standard	EKMA Ozone Estimate (ppb) City-Specific
1 August	203	26	26.6	55	18	1600	70	94	96
2 August	435	58	27.4	68	14	1700	80	158	131
3 August	701	83	28.6	90	19	1600	145	211	162
4 August	707	101	28.9	100	14	1300	147	222	168
5 August	387	64	31.3	82	14	1200	113	152	126
6 August	442	44	28.0	71	24	1600	105	147	126
7 August	285	35	26.1	66	24	1600	86	117	108
8 August	348	35	27.5	90	25	1800	124	126	114
11 August	781	96	33.2	82	13	1400	117	229	173
12 August	622	65	35.3	68	17	1400	166	188	149
13 August	728	86	33.7	109	9	1200	180	216	166
14 August	265	30	31.5	84	19	1400	160	109	104
16 August	484	58	28.4	68	21	1400	93	165	136
17 August	819	76	28.5	76	25	1700	148	217	167
18 August	967	100	29.2	92	25	1600	168	252	187
19 August	808	96	30.6	85	25	1600	133	233	175
20 August	1320	141	32.3	95	21	1500	127	314	221
21 August	1324	114	33.7	105	14	1500	169	293	209
22 August	315	43	34.6	89	18	1300	128	128	114
23 August	1432	159	33.5	88	20	1200	156	336	235
24 August	1704	127	33.3	112	20	1200	176	329	224
25 August	1330	158	33.1	111	15	1400	192	325	228
29 August	321	44	29.6	64	10	1400	75	130	115
30 August	542	55	28.9	68	22	1500	109	170	139
31 August	833	112	30.0	75	24	1300	130	245	181

TABLE 3 (continued)

Date (1976)	Precursors NMOC (ppbC)	Precursors NO_x (ppb)	Maximum Temperature (°C)	Background O_3 (ppb)	Observed Maximum Ozone Station	Observed Maximum Ozone Time (CDT)	Observed Maximum Ozone Concentration (ppb)	EKMA Ozone Estimate (ppb) Standard	EKMA Ozone Estimate (ppb) City Specific
1 Sept	1098	105	29.1	104	18	1200	134	268	195
2 Sept	529	58	29.7	74	21	1600	109	171	139
3 Sept	577	61	27.2	80	22	1500	126	179	144
4 Sept	592	27	35.0	67	18	1400	149	132	117
5 Sept	639	50	29.8	59	18	1500	111	173	142
6 Sept	483	42	29.9	82	18	1700	107	149	127
7 Sept	697	84	31.5	72	18	1400	114	211	162
8 Sept	895	65	31.4	63	21	1300	107	211	163
10 Sept	737	102	25.2	52	10	1300	58	227	170
11 Sept	735	78	26.9	60	22	1400	99	211	163
13 Sept	900	90	30.3	41	22	1400	117	237	178
14 Sept	1076	103	31.4	82	14	1600	118	264	193
15 Sept	1010	106	30.2	78	18	1600	107	261	191
16 Sept	454	63	27.3	68	24	1600	94	164	134
17 Sept	2055	189	28.6	70	12	1300	188	403	267
18 Sept	1267	118	31.5	74	14	1400	132	292	209
19 Sept	559	42	31.9	54	14	1300	91	156	132
23 Sept	717	97	24.4	53	18	1400	75	222	168
24 Sept	424	65	27.2	58	25	1600	106	160	131
25 Sept	694	86	25.6	56	18	1600	80	212	163
26 Sept	659	41	27.9	52	18	1200	73	162	134
30 Sept	548	93	24.7	61	23	1500	102	192	148
1 Oct	1884	210	31.5	78	2	1500	244	405	272
2 Oct	2858	302	33.4	81	15	1700	186	528	333
3 Oct	1088	150	30.5	105	14	1300	125	295	210

TABLE 3 (concluded)

Date (1976)	Precursors NMOC (ppbC)	Precursors NO_x (ppb)	Maximum Temperature (°C)	Background O_3 (ppb)	Observed Maximum Ozone Station	Observed Maximum Ozone Time (CDT)	Observed Maximum Ozone Concentration (ppb)	EKMA Ozone Estimate Standard	EKMA Ozone Estimate City-Specific
4 Oct	593	73	29.1	83	18	1300	104	191	150
12 Oct	890	124	29.7	73	22	1300	78	258	189
13 Oct	686	97	29.7	51	18	1500	58	217	164

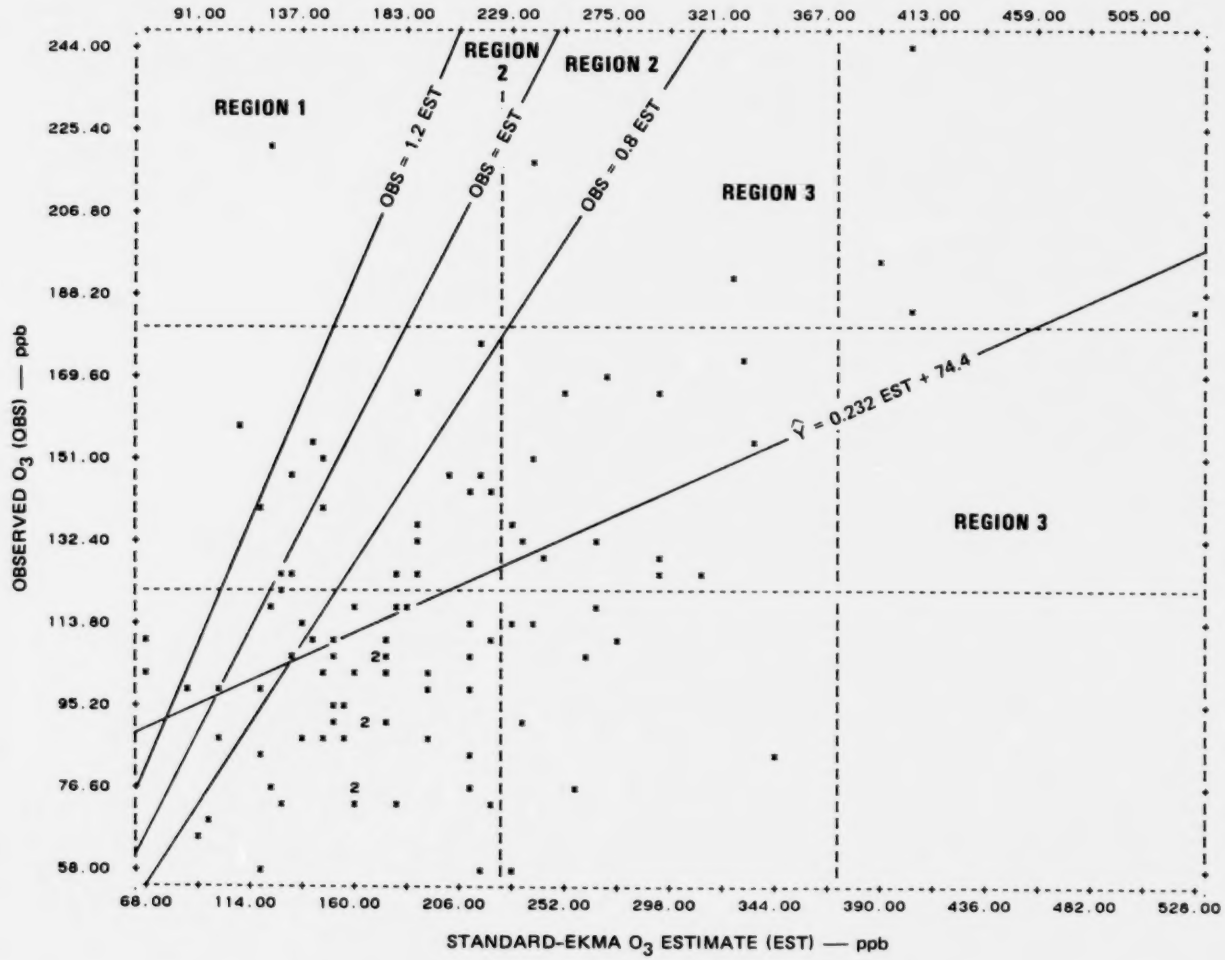


Figure 4. Scatterplot of observed ozone and standard-EKMA ozone estimate for St. Louis. Number of points plotted is 100.

and $\hat{Y}$ represents the value of OBS obtained from the regression line. The tendency of EST to overestimate OBS is reflected in the value of the slope of the regression line. The regression line implies that, on the average, $\hat{Y}/EST > 1.2$ for EST < 77 ppb, and $\hat{Y}/EST \leq 1.2$ otherwise. However, the regression line serves only as a rough guide to the location of a point in one of the three regions of Figure 4: We want to know the probability that a point is located in a particular region as a function of NMOC and NO_x. This will be examined in connection with the definition of the accuracy regions of the NMOC-NO_x plane.

The four points located in Region 1 (see Figure 4) are distinguished by having very low concentrations of NMOC or NO_x. Table 4 shows the 0600-0900 NMOC and NO_x levels at the individual monitoring sites and the corresponding spatial averages. (The spatial averages also appear in Table 3). For all the days in the evaluation data set, the spatially averaged NMOC range from 146 to 2858 ppbC and have a median concentration of 577 ppbC. Ranked in order of increasing concentration, the four NMOC spatial averages in Table 4 were lowest (146), second lowest (147), sixth lowest (228), and eighth lowest (265), where the number in parentheses is the NMOC concentration. NO_x ranged from 13 to 302 ppb, with a median level of 65 ppb. The NO_x spatial averages in Table 4 ranked lowest (13), second lowest (14), ninth lowest (30), and 27th lowest (45). The levels of both NMOC and NO_x for 12 and 24 July were the lowest in the evaluation data set. Sites 101 and 107 are solely responsible for the low NO_x on these two days; the NO_x data are missing for the other two sites. This raises the possibility that Sites 101 and 107 may not be representative of the areawide NO_x level on these two days; the same argument applies to NMOC.

Table 4 shows that on 13 July the EKMA estimate underpredicted the observed value by a large margin. In this case the problem appears to be the NMOC rather than the NO_x. As Table 4 shows, Site 101 reported a very low NMOC of 41 ppbC while the NO_x was 48 ppb, for an NMOC/NO_x ratio of 0.85. [This low ratio does not violate our data selection criterion (NMOC/NO_x > 5) because the criterion pertains to the spatially averaged NMOC and NO_x, not to the individual monitoring sites.] Such a low ratio is unusual and suggests that this NMOC measurement at Site 101 may be incorrect. The low value of NMOC causes the spatial average to be low, which in turn produces a low ozone estimate. This situation recurs on 14 August, when Table 4 shows that Site 105 reported a relatively high value of NMOC, but that Site 106 has a very low value. The latter site also has a very low NMOC/NO_x ratio.

All four cases of substantial underprediction are associated with data of questionable reliability; but, because vagaries in the data will always occur, these four cases will be retained and used in subsequent analyses.

Relationships Between the Variables

Table 5 shows the correlation coefficients for selected variables in the St. Louis evaluation data set. NMOC and NO_x are highly correlated; the implications of this relationship will be discussed below. The table shows that EST is highly correlated with NMOC and NO_x. This is expected, because it reflects the strong correlation between NMOC and NO_x. However, the

TABLE 4. SELECTED 0600-0900 NMOC AND NO_x CONCENTRATIONS*
AT SOURCE-REGION MONITORING SITES IN ST. LOUIS

Date	Monitoring Site								Spatial Average		Ozone (ppb)	
	101		105		106		107					
	NMOC	NO_x	NMOC	NO_x	NMOC	NO_x	NMOC	NO_x	NMOC	NO_x	Observed	Standard EKMA Estimate
12 Jul	115	13	270	---	---	--	53	13	146	13	105	68
13 Jul	41	48	---	---	273	--	371	43	228	45	223	121
24 Jul	50	14	---	---	---	--	245	15	147	14	112	70
14 Aug	---	--	501	---	29	23	---	38	265	30	160	109

*NMOC in ppbC NO_x and ozone in ppb. Dashes denote missing data.

TABLE 5. MEANS, STANDARD DEVIATIONS AND CORRELATIONS
FOR SELECTED VARIABLES IN THE ST. LOUIS EVALUATION DATA SET

Variable	Mean	Standard Deviation	Correlation Coefficient*							
			NO_x	1/NMOC	$1/NO_x$	Maximum Temperature	Background O_3	OBS	EST†	OBS/EST
NMOC (ppbC)	690.6	448.1	0.93	-0.72	-0.63	--	0.27	0.54	0.98	-0.49
NO_x (ppb)	76.5	46.1		-0.69	-0.70	--	0.23	0.44	0.97	-0.55
$1/NMOC$ ($ppbC^{-1}$)	0.0020	0.0013			0.90	--	-0.19	-0.31	-0.79	0.72
$1/NO_x$ (ppb^{-1})	0.0180	0.0118				--	-0.20	-0.26	-0.74	0.73
Maximum temperature (^oC)	30.4	3.2					0.30	0.47	0.09	0.35
Background O_3 (ppb)	78.0	19.6						0.63	0.27	0.22
OBS (ppb)	119.9	37.0							0.49	0.32
EST (ppb)†	196.0	78.4								-0.58
OBS/EST	0.676	0.280								1.0

*All correlations significant at the 0.05 level or better. Dashes denote that correlation is not statistically significant.

†Standard-EKMA estimate.

correlation between OBS and NMOC and OBS and NO_x is not high; the coefficients are 0.54 and 0.44, respectively. The variables $1/NO_x$ and 1/NMOC correlate best with the ratio OBS/EST, the former being slightly better. Temperature correlates best with observed O_3, which reflects the indications of Figure 2. There is no statistically significant correlation between temperature and NMOC, NO_x, their respective inverses, and EST.

The scatterplot of NO_x and NMOC concentrations (Figure 5) shows the combinations of these precursors represented in the evaluation data set. The shaded region of the scatterplot is excluded from consideration by the selection criterion of $NMOC/NO_x$ greater than 5.0. Figure 5 provides graphic evidence of the high correlation between NO_x and NMOC ($r = 0.93$).

The strong linear relationship between NMOC and NO_x indicates that the source-region sites are influenced by sources with a relatively uniform $NMOC/NO_x$ ratio. Such uniformity suggests that area sources, mainly mobile sources, are the primary contributors of NMOC and NO_x in the area of central St. Louis during 0600-0900.

As is evident in Figure 5, the interdependence of NMOC and NO_x causes the points to fall along a narrow band. This limits the region of the EKMA diagram in which the evaluation can be performed. In this region, NMOC varies from 146 to 2,858 ppbC and NO_x from 13 to 302 ppb. This region may be defined by the 95-percent confidence band for the linear regression of NO_x on NMOC. The fitted regression equation is

$$NO_x = 10.4 + 0.0958(NMOC),$$

where NO_x is in ppb, NMOC is in ppbC, and the standard error of estimate is $s = 16.8$. The upper and lower boundaries of the region are defined by the lines $NO_x \pm 1.96s$. In what follows, we will refer to the region bounded by these lines as the evaluation region. (To be technically correct, the confidence limits drawn on the scatterplot should be concave arcs rather than straight lines; however, the amount of curvature is small and the straight-line approximation is satisfactory for this application.)

As Table 5 indicates, the OBS/EST ratio correlates best with the variable $1/NO_x$. The scatterplot of OBS/EST and $1/NO_x$ (Figure 6) shows a nicely correlated linear relationship ($r = 0.730$) that is well behaved, because the residual variance about the fitted regression line appears to be constant. The fitted regression line is $OBS/EST = 0.365 + 17.3(1/NO_x)$, where NO_x is in ppb, and the standard error of estimate is $s = 0.193$. This regression line is plotted in Figure 6.

Because NO_x and NMOC are correlated, the relationship between OBS/EST and 1/NMOC is similar to that with $1/NO_x$, but, as Table 5 shows, the correlation coefficient is slightly smaller for 1/NMOC. These dependencies will be exploited in the next section to derive a multiple regression equation for OBS/EST as a function of several variables.

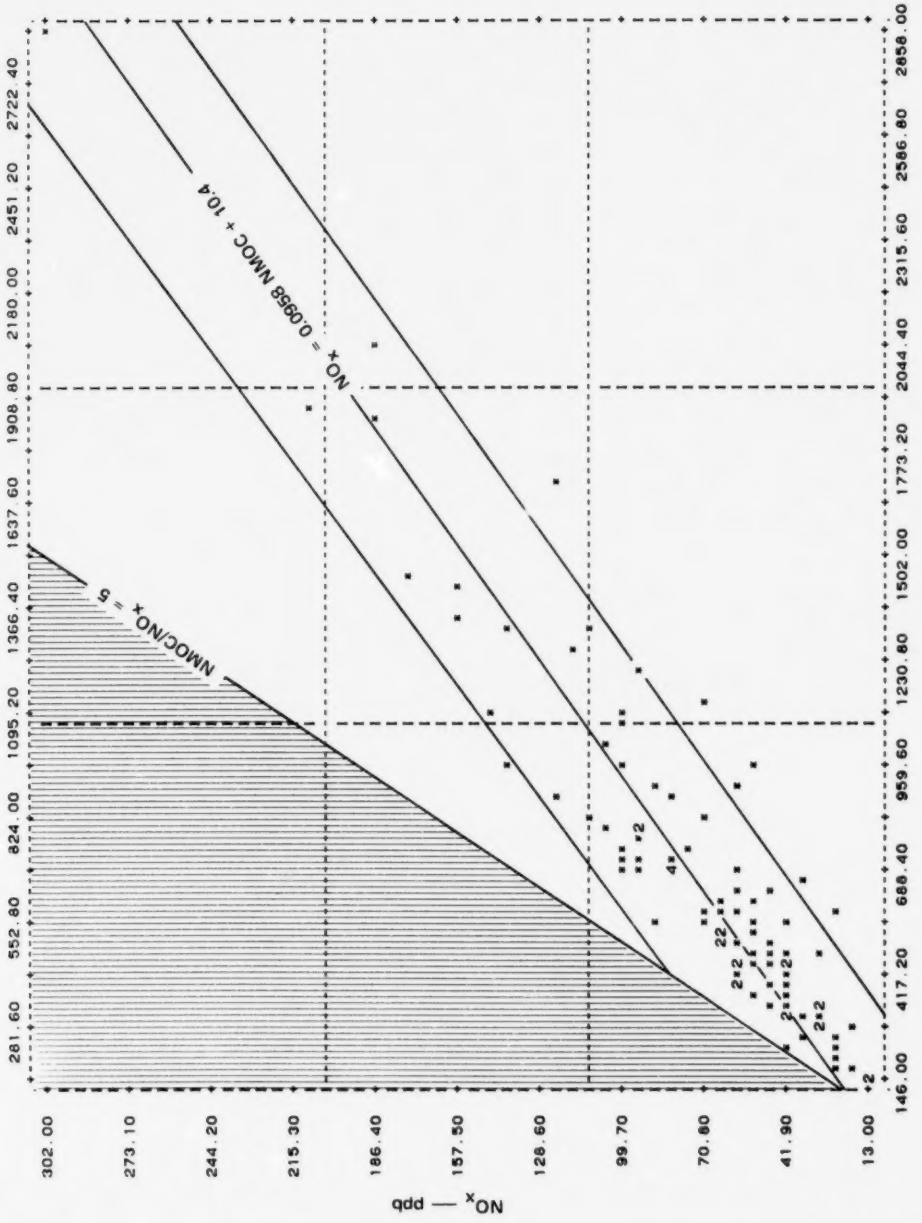


Figure 5. Scatterplot of 0600–0900 (CDT) concentrations of NMOC and NO_x for St. Louis.

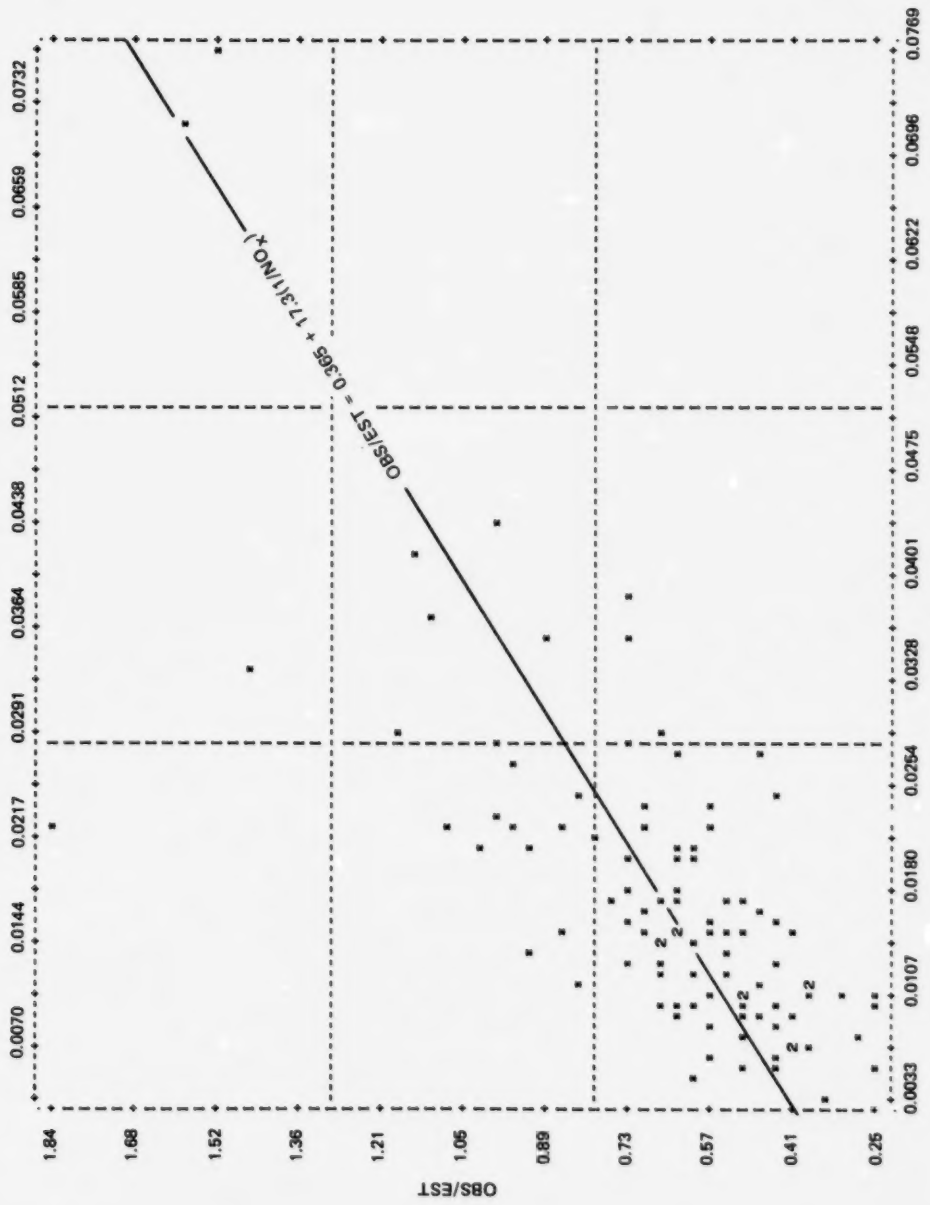


Figure 6. Scatterplot of OBS/EST as a function of $1/NO_x$ for standard-EKMA ozone estimate in St. Louis.

Definition of Accuracy Regions

To define the accuracy regions, and associated probabilities, in the O_3 isopleth diagram, a multiple regression equation was derived for the OBS/EST ratio as a linear function of $1/NO_x$, 1/NMOC, daily maximum temperature (T), and background ozone (BO3). The equation is:

$$OBS/EST = -0.628 + 8.68(1/NO_x) + 99.23(1/NMOC)$$

$$+ 0.00451(BO3) + 0.0196(T) \quad , \qquad (2)$$

where NO_x and BO3 are in ppb, NMOC in ppbC, and T in ^{o}C. All the coefficients of the regression are statistically significant ($p < 0.005$); the multiple regression coefficient is 0.86, and the standard error of estimate is $s = 0.146$.

Using Eq. 2, we can estimate the probability that OBS/EST > 1.2, or $0.8 \leq$ OBS/EST ≤ 1.2, or OBS/EST < 0.8 as a function of NMOC and NO_x for any specified values of BO3 and T. This is illustrated below.

Following the notation of Section Two, define R = OBS/EST, $R_1 = 1.2$, $R_2 = 0.8$; denote Eq. 2 by $R = F(NO_x, NMOC, BO3, T)$; and let $P(R \leq R_j)$ = probability that $R \leq R_j$, $j = 1,2$. Then

$$P(R \leq R_j) = \Phi \frac{[R_j - F(NO_x,NMOC,BO3,T)]}{s} \quad , \qquad (3)$$

where Φ is the cumulative normal probability distribution function and $s = 0.146$, the standard error of estimate of Eq. 2. Then we can use Eq. 1 in conjunction with Eq. 3 to calculate the probability that the ratio R = OBS/EST is located in one of the three accuracy regions, i.e. $P(R > 1.2)$, $P(0.8 \leq R \leq 1.2)$, and $P(R < 0.8)$. The procedure is simple: First, fix BO3 and T by, for example, setting them equal to their mean values, which were defined in Table 5. Thus, BO3 = 78.0 ppb, and T = 30.4°C. Substituting these values in Eq. 2 yields

$$F(NO_x,NMOC,78.0,30.4) = 0.3196 + 8.68/NO_x + 99.23/(NMOC) \quad . \qquad (4)$$

Equation 4 is then substituted in Eq. 3 to obtain the desired probabilities.

Figure 7 shows a plot of the probability as a function of a new variable $Z = 8.68/NO_x + 99.23/NMOC$. Using Figure 7, one can calculate the probability associated with each accuracy region for a given (NMOC, NO_x) pair. In general, Figure 7 indicates that there is a high probability of overprediction [i.e. $P(R<0.8)$] when NMOC and NO_x are large (which corresponds to low values of Z). It is also apparent from Figure 7 that there is a high probability of underprediction [i.e. $P(R > 1.2)$] for low values of NMOC and NO_x. Relatively accurate predictions occur in the interval $0.56 \leq Z \leq 0.80$, where $P(0.8 \leq R \leq 1.2)$ varies between 0.70 and 0.83, the maximum probability of 0.83 occurring in the neighborhood of Z = 0.68.

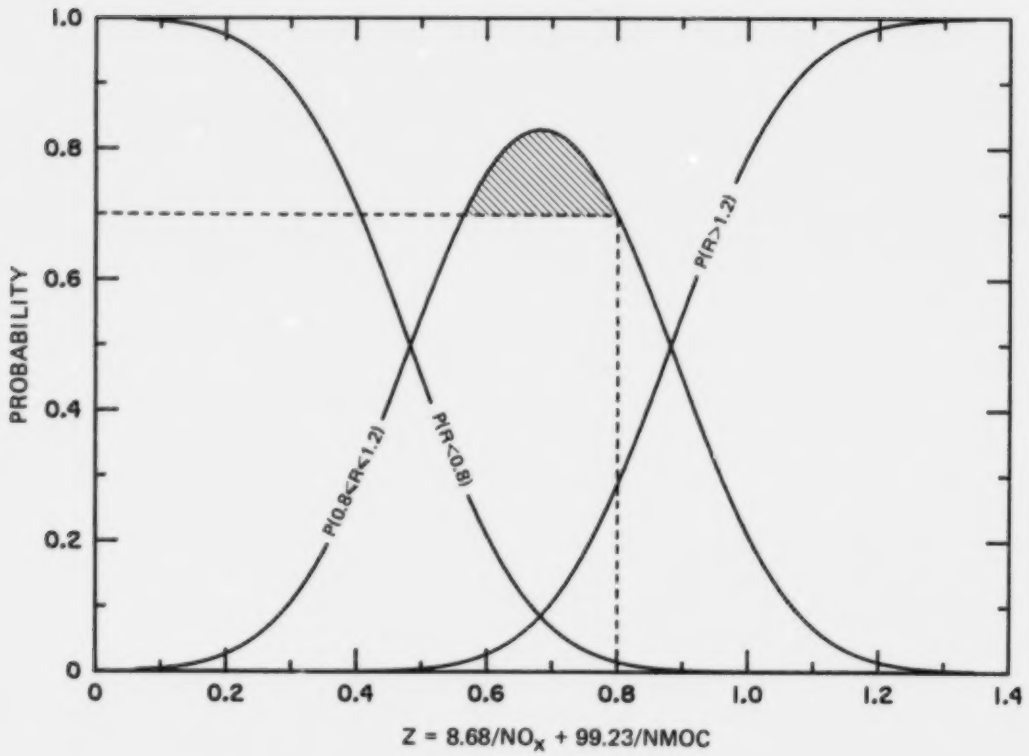


Figure 7. Accuracy probability plot for standard-EKMA ozone estimates for St. Louis. Mean values are assumed for background ozone and maximum daily temperature.

Given specific values of NMOC and NO_x, it is easy to compute the corresponding value of Z to find the probability in Figure 7. Figure 8 complements Figure 7 by showing the values of NMOC and NO_x that correspond to selected values of Z. Figure 8 also displays the evaluation region for NMOC and NO_x that was previously defined in Figure 5. Each constant-Z curve in Figure 8 can be related to the probability plots of Figure 7; for example, the shaded area in Figure 7, which corresponds to the interval $0.56 \leq Z \leq 0.80$, maps into the shaded area of Figure 8, which defines the values of NMOC and NO_x that are associated with the most accurate predictions. High probabilities of overprediction, where $P(R < 0.8) > 0.70$, occur for $Z \leq 0.3$; these curves are seen in Figure 8 to be linked to high values of NMOC and NO_x. A high probability of underprediction occurs for $Z \geq 1.0$; Figure 8 shows that this is a very small area within the evaluation region.

In view of the above, the standard-EKMA ozone estimate can be considered to be an upper bound for the observed daily maximum in the region defined by $Z \leq 0.8$ in Figures 7 and 8. In this region, the probability that $R \leq 1.2$ is approximately 0.7 at $Z = 0.8$ and increases rapidly for $Z < 0.8$.

The effect of changing B03 and T is shown in Figure 9, which contains probability plots similar to those in Figure 7 but with B03 and T set to their respective mean value plus two standard deviations, viz., B03 = 117 ppb and T = 37°C. Comparison of Figures 7 and 9 shows that the higher B03 and T have shifted the location of the probability curves to the left, the shift causing the curves for $P(R < 0.8)$ and $P(0.8 \leq R \leq 1.2)$ to be truncated for $Z \leq 0.1$. The new location of the probability curves implies that accurate predictions are most likely in the interval $0.25 \leq Z \leq 0.49$, with the maximum located approximately at $Z = 0.37$. Referring to Figure 8, it can be seen that within the evaluation region the area corresponding to $0.25 \leq Z \leq 0.49$ is larger than the shaded area. However, in Figure 9 underprediction is most likely to occur for $Z \geq 0.65$, and the corresponding area in Figure 8 is larger than the underprediction region associated with Figure 7. In Figure 9, the standard-EKMA ozone estimate can be considered an upper bound of the observed daily maximum for $Z \leq 0.49$, whereas in Figure 7 it was for $Z \leq 0.8$.

In general, increasing either B03 or T, or both, will shift the probability curves to the left. This lowers the probability of overprediction, expands the area of highest accuracy, and increases the area where the likelihood of underprediction is high. Lowering B03 or T shifts the probability curves to the right, thus reversing the effect associated with the leftward shift.

EVALUATION OF CITY-SPECIFIC EKMA

City-Specific Ozone Estimates

The EKMA computer program was run using the input listed in Table 6 to generate city-specific (C-S)O_3 estimates for St. Louis. Table 3 lists the values of the C-S ozone estimates, hereafter denoted by CS03. The date shown in Table 6 was used to calculate CS03 because the highest observed ozone level in the evaluation data set occurred on that date (cf. Table 3). The inversion

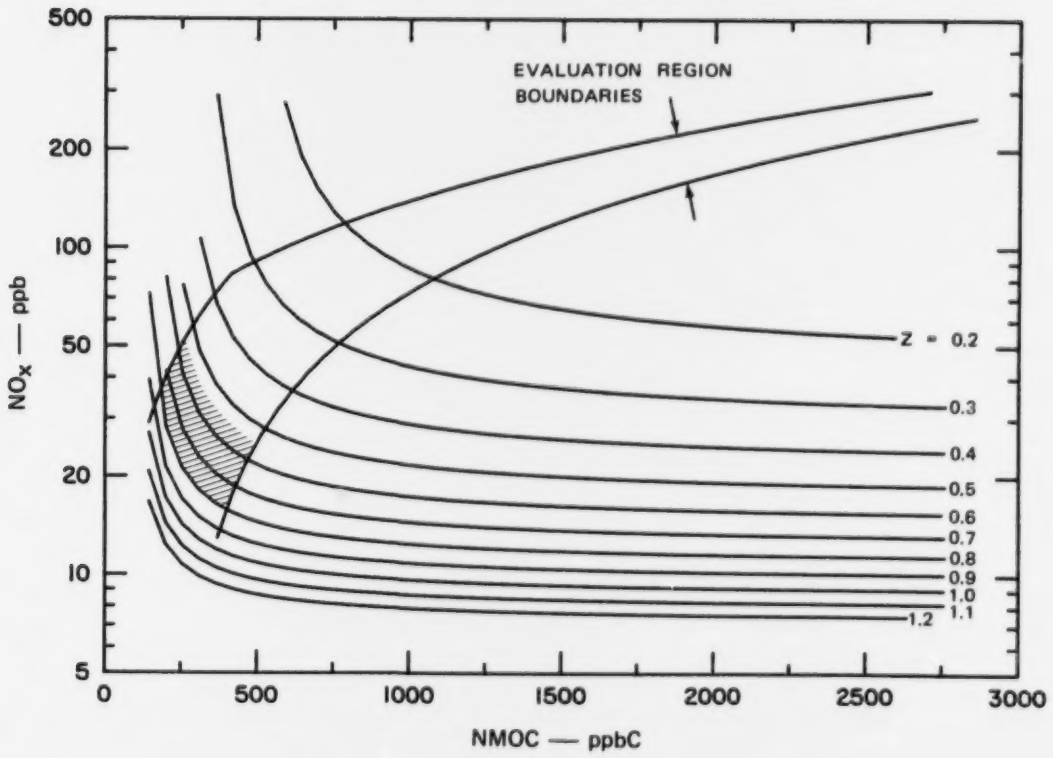


Figure 8. Plot of constant-Z curves on NMOC-NO_x plane for standard-EKMA estimates for St. Louis. Shading denotes area where ozone estimates are most accurate, assuming average conditions for temperature and background ozone.

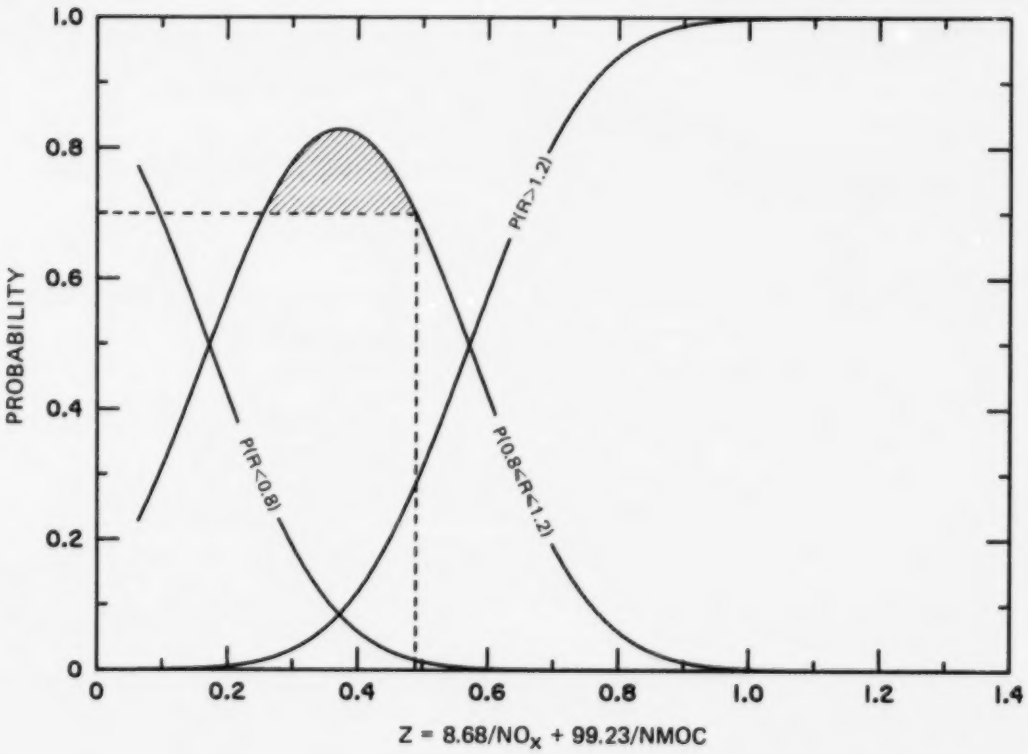


Figure 9. Accuracy probability curves for standard-EKMA evaluation for St. Louis. Background ozone equals 117 ppb and maximum temperature equals 37° C.

TABLE 6. SUMMARY OF INPUT PARAMETERS FOR OBTAINING CITY-SPECIFIC EKMA OZONE ESTIMATES FOR ST. LOUIS

Parameter	Value
Date	1 October 1976
Location	Latitude 38.4°N, Longitude 90.15°W
Inversion height data	
Initial height (m)	400
Final height (m)	1500
Start time of rise	0900
Ending time of rise	1300
Post-0800 hourly NO_x and NMOC hourly emissions fractions by ending hour	
0900	0.24
1000	0.21
1100	0.09
1200	0.09
1300	0.06
1400	0.03
Reactivity, NO_2/NO_x	0.59
Background ozone concentration (ppm)	0.092

height data represent an average estimated from several days of temperature soundings. Post-0800 emissions fractions are average values calculated from trajectories for the five days with the highest observed ozone in the evaluation data set. The /NO_x ratio and the background ozone level are the mean values of these quantities for all the days in Table 3 for which the observed ozone exceeded 120 ppb.

The C-S and standard-EKMA ozone estimates are highly correlated. Figure 10 shows a scatterplot of CS03 and EST, where EST denotes the standard ozone estimate. The correlation coefficient is r = 0.99; the regression line is CS03 = 0.563(EST) + 42.8, where the standard error of the regression is s = 1.65, and the units are ppb. The fitted line implies that the ratio CS03/EST < 1 for EST $\geq$ 98 ppb and that CS03/EST > 1 for EST < 98 ppb. The presence of a statistically significant intercept in the regression line indicates that the location of the C-S isopleths has been shifted with respect to the standard-EKMA isopleths.

Comparison of Observed and City-Specific Ozone Estimate

Figure 11 shows a scatterplot of OBS as a function of CS03. Of the 100 cases, eight are in Region 1, 36 are in Region 2, and 56 are in Region 3. By contrast, 4 percent of the standard-EKMA estimates were located in Region 1, 18 percent in Region 2, and 78 percent in Region 3. Thus, although the C-S estimates tend to underpredict the observations more frequently, the fraction of cases in Region 2 for the C-S method is twice that for the standard EKMA. Overall, the percentages of cases satisfying the inequalities OBS/EST $\leq$ 1.2 EST and OBS/CS03 $\leq$ 1.2 were 96 and 92 percent, respectively, which is a relatively small difference. Hence, the C-S estimates also produce an upper bound for the observed ozone, but more cases are underestimated.

As was the case with EST, OBS and CS03 are correlated. The correlation coefficient is r = 0.49, and is highly statistically significant. The regression line relating OBS and CS03 is plotted in Figure 11, and is defined by OBS = 0.407(CS03) + 32.5, where the units are ppb, and the standard error of the regression is s = 32.5 ppb. Figure 11 shows that the estimates obtained from the regression equation always are in Regions 2 and 3 for CS03 $\geq$ 83 ppb.

The eight points in Region 1 of Figure 11 include the four points in Region 1 of Figure 4 and four additional points for 8 June, 8 and 29 July, and 4 September. Two of the points in Region 1 of Figure 11 appear to have NO_x levels that are considerably lower than would be expected for the concurrent NMOC concentration. One of the two points in question is 8 June, which has NMOC = 1,111 ppbC and NO_x = 75 ppb. The NMOC level is one of the highest in the evaluation data set, ranking 87th out of 100, where the rank increases with increasing concentration. By contrast, 75 ppb ranked 60th among all the NO_x values, and is more than two standard errors smaller than the NO_x estimated from the regression equation relating NO_x and NMOC (see Figure 5). The latter is also true for 4 September, which has NMOC = 592 ppbC and NO_x = 27 ppb. In this case, NMOC ranked 52nd, but NO_x ranked sixth. The NMOC and NO_x levels for the remaining two points (for 8 and 29 July) appear to be consistent with each other.

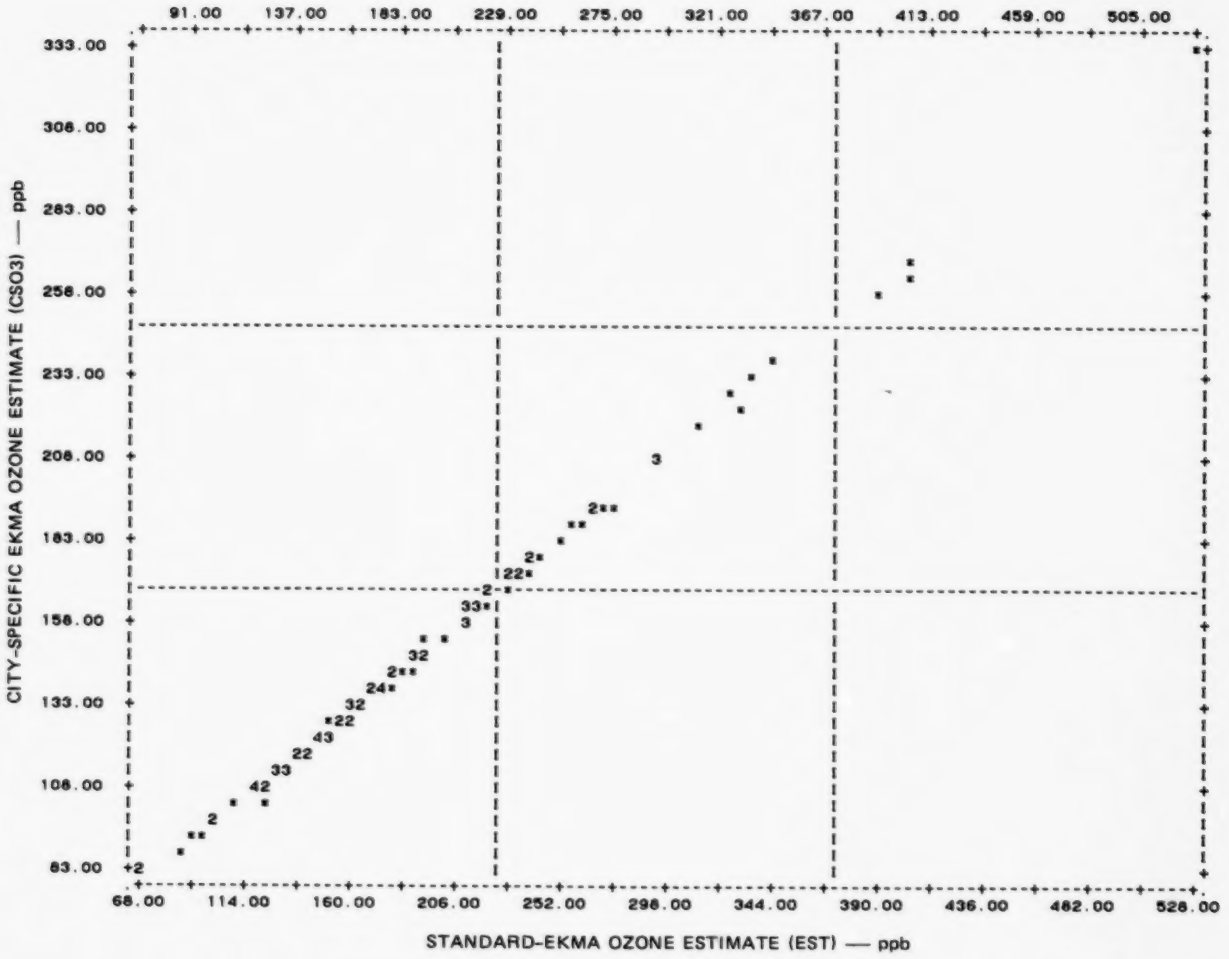


Figure 10. Scatterplot of city-specific and standard EKMA ozone estimates for St. Louis.

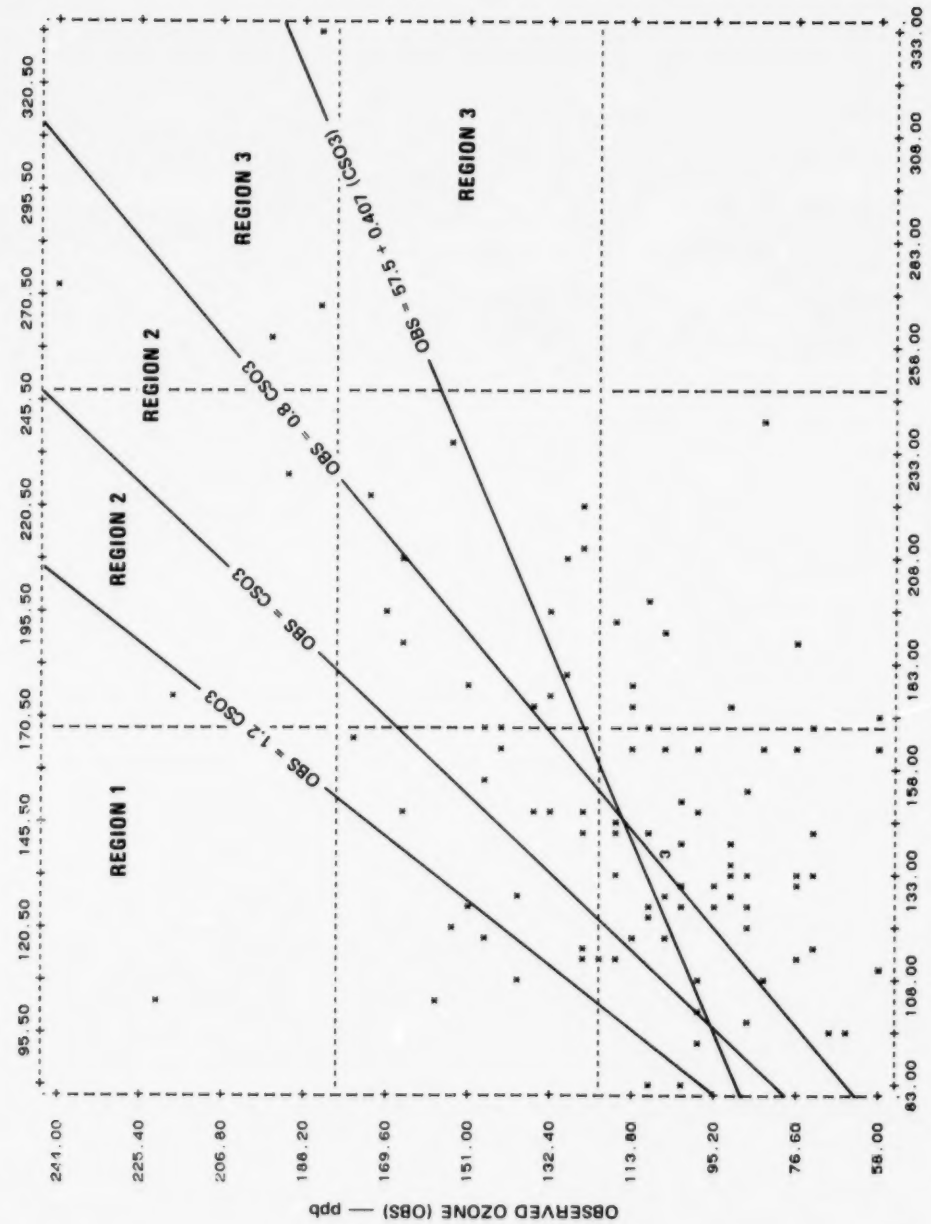


Figure 11. Scatterplot of observed ozone and city-specific EKMA ozone estimate for St. Louis.

Relationships Between the Variables

Table 7 shows the means, standard deviations and pairwise correlations for the variables of interest; the table is abbreviated because some of the information was previously presented in Table 5. As with the standard-EKMA ozone estimate, CSO3 is strongly correlated with NMOC and NO_x. The table also shows that CSO3 is correlated with background O_3, although the correlation is low. The ratio OBS/CSO3 correlates best and about equally well with $1/NO_x$ and 1/NMOC, followed by maximum daily temperature.

Definition of Accuracy Regions

A multiple regression equation was derived for OBS/CSO3 as a function of 1/NMOC, daily maximum temperature (T), and background O_3(BO3). The equation is

$$\text{OBS/CSO3} = -0.584 + 118.39(1/\text{NMOC}) + 0.0241(T) + 0.00543(\text{BO3}) \quad , \qquad (5)$$

where the units are as defined in Table 7. All the regression coefficients are statistically significant ($p < 0.002$), the multiple regression coefficient is 0.74, and the standard error of the regression is 0.182. In contrast to Eq. 2 for OBS/EST, the variable $1/NO_x$ does not appear in the regression. The stepwise regression procedure attempted to include $1/NO_x$, but its coefficient was not significant at the 0.05 level; NO_x was also tried with the same result. Equation 5 also has a smaller multiple regression coefficient, 0.74, compared to 0.86 for Eq. 2. Hence, Eq. 5 explains a smaller fraction of the variance than does Eq. 2. Moreover, the standard error for Eq. 5 ($s = 0.182$) is greater than that for Eq. 2 ($s = 0.146$).

Using Eq. 5, we plotted the accuracy probabilities as a function of $Z = 118.39/\text{NMOC}$ after setting $T = 30.4^{\circ}C$ and BO3 = 78.0 ppb, which are their respective means; the plot is shown in Figure 12. Complementing Figure 12, a family of lines for several values of Z is displayed in Figure 13, which also shows the evaluation region in the NMOC-NO_x plane.

Figure 12 shows that there is at least a 75 percent probability of satisfying the inequality $R \leq 1.2$ for $Z \leq 0.50$, which is the area to the left of the dashed vertical line shown in Figure 12. The area of Figure 12 in which $Z \leq 0.50$ corresponds to the region of Figure 13 in which NMOC $\geq$ 237 ppbC; the left boundary of the shaded area in Figure 13 is the line NMOC = 237 ppbC. Thus, inside the evaluation region of Figure 13, (NMOC, NO_x) combinations within and to the right of the shaded area have at least a 75-percent probability of producing an ozone estimate that is an upper bound for the observed daily maximum ozone.

The shaded area of Figure 12 is the region where the probability $P(0.8 \leq R \leq 1.2)$ attains its highest values, hence it is associated with ozone estimates that are most likely to be accurate. Within the shaded area, the variable Z varies from 0.35 to 0.50, and $P(0.8 \leq R \leq 1.2)$ varies from 0.70 to 0.73, reaching its maximum value of 0.73 in the neighborhood of Z = 0.43. The

TABLE 7. MEANS, STANDARD DEVIATIONS, AND CORRELATIONS FOR CITY-SPECIFIC EKMA EVALUATION FOR ST. LOUIS

Variable	Mean	Standard Deviation	Correlation*	
			CS03	OBS/CS03
NMOC (ppbC)	690.6	448.2	0.97	-0.34
NO_x (ppb)	76.5	46.1	0.97	-0.41
$1/NMOC$ ($ppbC^{-1}$)	0.0020	0.0013	-0.79	0.49
$1/NO_x$ (ppb^{-1})	0.0180	0.0118	-0.74	0.49
Maximum temperature (^oC)	30.4	3.2	--	0.41
Background O_3 (ppb)	78.0	19.6	0.26	0.38
OBS (ppb)	119.9	37.0	0.49	0.55
CS03 (ppb)	153.2	44.1	1.0	-0.42
OBS/CS03	0.814	0.266	-0.42	1.0

*Significant at the 0.05 level or better. Dashes indicate correlation is not statistically significant.

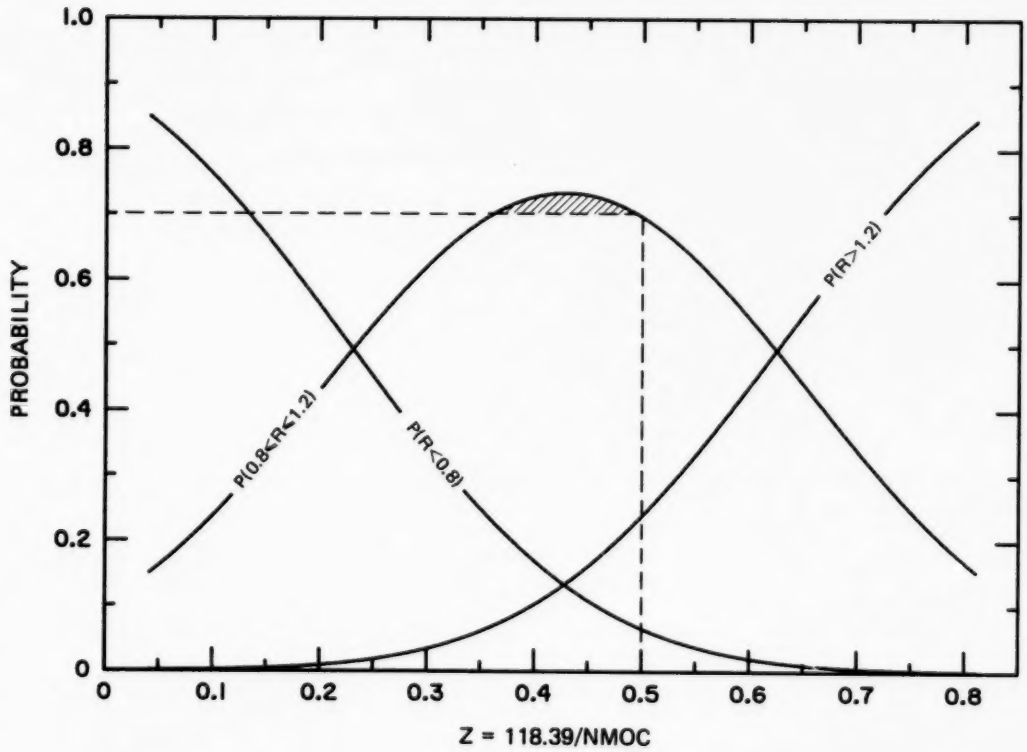


Figure 12. Accuracy probability plot for city-specific EKMA ozone estimates for St. Louis. Mean values are assumed for background ozone and maximum daily temperatures.

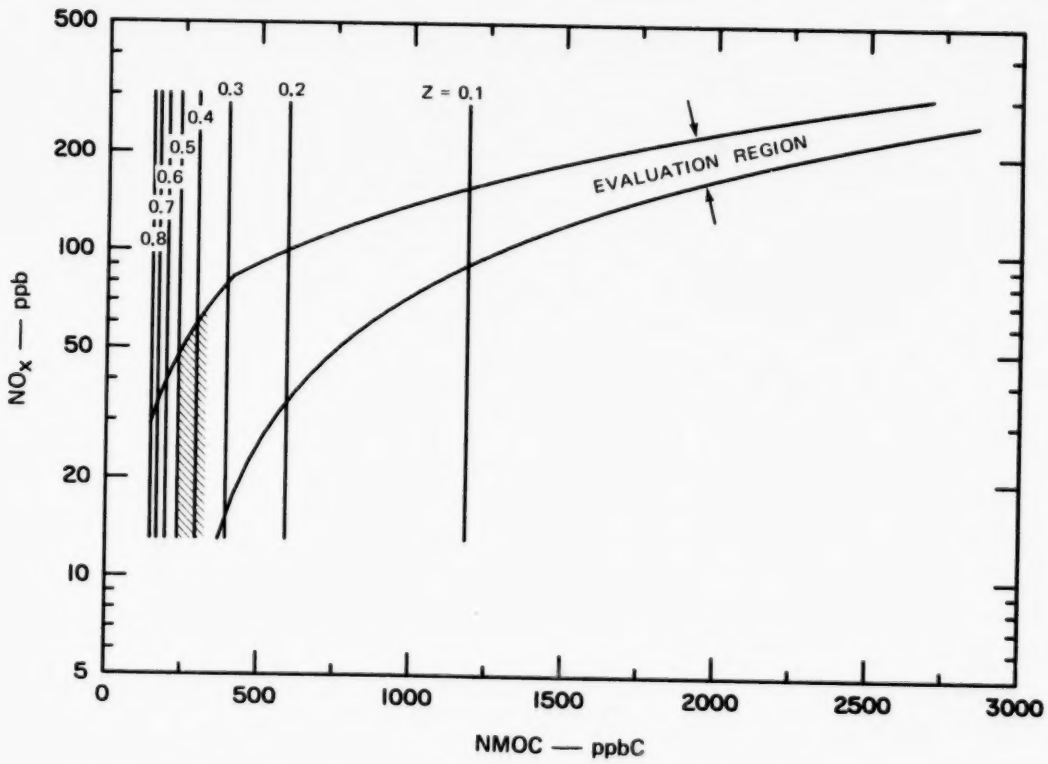


Figure 13. Plot of constant-Z lines on NMOC-NO_x plane for city-specific EKMA estimates for St. Louis. Shading denotes area where ozone estimates tend to be most accurate, assuming average conditions for maximum temperature and background ozone.

shaded area of Figure 13 corresponds to that of Figure 12. Inside the shaded area of Figure 13, NMOC is defined by $237 \leq$ NMOC ≤ 338 ppbC, and NO_x is limited by the boundaries of the evaluation region. Thus, (NMOC, NO_x) combinations in the shaded area of Figure 13 are associated with the C-S ozone estimates most likely to be accurate.

The maximum value of $P(0.8 \leq R \leq 1.2)$ in Figure 7 is higher than that in Figure 12--the respective maxima being 0.83 and 0.73--largely because the standard deviation for Eq. 2 is smaller than that for Eq. 5 ($s = 0.146$ and $s = 0.182$, respectively). The smaller standard deviation yields a narrower and taller probability curve for Eq. 2 (cf. Figure 7) in comparison with that of Eq. 5 (see Figure 12).

Comparison of Figures 8 and 13 shows that the location of the shaded area has been shifted upward in Figure 13. As a result, the values of NMOC and NO_x associated with the ozone estimates most likely to be accurate are lower for the standard EKMA (Figure 8) than for the C-S case (Figure 13). The relative placement of the shaded area in Figures 8 and 13 is also indicative of the difference in the predictive performance of standard and C-S EKMA. Thus, overprediction is more common for the standard than for C-S EKMA, as reflected by the size of the region above the shaded area in Figure 8. This situation is reversed in the case of underprediction; the region below the shaded area is smaller in Figure 8 than in Figure 13.

DISCUSSION

The results presented above indicate that both standard and C-S EKMA can be used to obtain ozone estimates that are upper bounds for the observations, where the upper bound criterion is observed $O_3 \leq 1.2$ estimated O_3. More precisely, an estimate has a high probability of being an upper bound if its (NMOC, NO_x) coordinates fall in certain specific areas of the NMOC-NO_x plane (see Figures 8 and 13). In general, high NMOC and NO_x lead to upper bound estimates, and low NMOC and NO_x to underestimates, in which observed $O_3 > 1.2$ times estimated O_3.

The analysis showed that within certain bounds, the standard EKMA can be used to provide realistic estimates of maximum ozone levels for the St. Louis area. Three regions of predictive accuracy were defined wherein the ratio of observed to estimated ozone is greater than 1.2 between 0.8 and 1.2, and less than 0.8. Of the 100 estimates, 78 fell in the third region, which is indicative of substantial overprediction (as expected). Eighteen estimates fell in the second region, where EKMA is accurate to within ±20 percent, indicating that reasonably credible estimates can be obtained with the standard EKMA. Four cases fell in the first region, where an underestimate occurs, which suggests that the standard EKMA could be used as a screening tool with the expectation that there is a low probability that its predictions would be exceeded.

The C-S EKMA's ozone estimates were generally more accurate than the standard EKMA's. Thus, although the C-S EKMA also produced upper bounds for the observations, 36 percent of the C-S estimates fell in the region where $0.8 \leq$ observed/estimated ≤ 1.2, compared to 18 percent for the standard-EKMA

estimates. However, the C-S EKMA displayed a greater tendency to underestimate the observations than did the standard EKMA (8 percent compared to 4 percent, respectively).

The accuracy regions for the standard-EKMA estimates were shown to be functions of NMOC, NO_x, background ozone, and maximum daily temperature. Thus, to define the accuracy regions on the NMOC-NO_x plane it is necessary to assign a value to background ozone and temperature. One of the two examples presented used mean values for these two quantities (cf. Figures 7 and 8); the other used the mean value plus two standard deviations (cf. Figure 9). These two examples demonstrated how the accuracy regions are defined and mapped on the NMOC-NO_x plane. They also showed the effect of changing the background ozone and temperature on the position and size of the accuracy regions. When applied to a particular problem, it is desirable to tailor the accuracy regions to the specific prevailing background ozone and temperature. Having defined the values of these two variables, it is a simple matter to use Eq. 2 and 3 to generate a plot similar to Figure 7. In so doing, care must be taken to use values of background ozone and temperature that fall within the range of values shown in Table 3, because these were used to derive Eq. 2. Similar considerations apply to the accuracy regions associated with the C-S estimates.

In discussing the probability $P(0.8 \leq R \leq 1.2)$ in connection with Figures 7, 9, and 12, we used 0.70 as the minimum probability that defines the region where the ozone estimates are most likely to achieve the highest accuracy. The value of 0.70 was selected because we desired to have a relatively high probability of being in the high-accuracy region, but not one so high that the size of the region would be minuscule. Examination of Figures 7, 9, and 12 suggested that 0.70 would be adequate. For another problem, it may be more appropriate to use some other probability threshold to define the boundaries of the high-accuracy region.

SECTION 4

EKMA EVALUATION FOR THE HOUSTON AREA

DATA REVIEW AND ANALYSIS

The data used were collected during the Houston Area Oxidant Study (HAOS) and the Houston Oxidant Modeling Study (HOMS). The HAOS data span the period May-October 1977; the HOMS data cover about one month, from 15 September through 12 October 1978. Comprehensive analyses and descriptions of these two data bases are found in the references given in Table 1.

Analysis of the Houston data did not reveal any strong linear correlations between maximum ozone and other pollutants or meteorological parameters. The low correlation between peak ozone concentration and daily maximum temperature can be explained by the consistently high daily maximum temperature observed in the Houston area during the period of interest. The best correlation ($r = 0.53$) was found between the maximum ozone concentration and the difference between the daily maximum temperature and the morning minimum temperature, where the latter is defined as the lowest temperature between midnight and 0800. A low but statistically significant correlation was found between maximum daily ozone with daily peak hourly solar radiation and with the 1000-1400 averaged solar radiation values for the HAOS data. The data also suggest that there is a minimum level of solar radiation that must exist for high ozone concentrations to occur (Maxwell, 1981).

Five monitoring sites, located at Aldine, Mae Drive, Crawford, Clinton, and Parkhurst, were selected as source-region monitors that define the 0600-0900 (CDT) NMOC and NO_x concentrations. The analysis that led to this selection was previously reported by Ludwig and Martinez (1979). Briefly, the analysis showed that these five sites have similar NMOC/NO_x ratios, and that mobile source emissions predominate during 0600-0900.

DEFINITION OF EVALUATION DATA SET

The review of the HAOS data showed that NMOC and NO_x measurements were very sparse during the period August-October. As a result, the HAOS evaluation data set is almost entirely composed of data for the period May-July

1977. Other data availability considerations also reduced the size of the HOMS evaluation data set. The criteria used to select Houston test days for evaluating EKMA are as follows:

- Maximum daily temperature at least 24°C.
- Difference between maximum and morning temperature greater than 6°C.
- 0600-0900 areawide average NO_x at least 36 ppb.
- Average 1000-1400 solar radiation at least 0.22 ly/min.

(Recall that the conditions represented by these criteria are always present when the daily maximum ozone is at least 100 ppb.) The fourth criterion was applied only to the HAOS data, because no solar radiation data were available for the HOMS. A few days were discarded because thunderstorms or rain showers preceded the time of the observed peak ozone concentration. Other days were rejected because the ozone maximum occurred before 1200. Missing values of the area-wide average 0600-0900 NO_x concentration resulted in more days being discarded than did any other criterion.

The HAOS and HOMS evaluation data sets are shown in Tables 8 and 9, respectively. The number of days in the HAOS data set is 61; the HOMS data set contains 17 days. (A complete evaluation of EKMA could not be performed using the HOMS data set because of its small size.)

The performance of the EKMA was investigated separately using the HAOS and HOMS data because the two data sets differ significantly in several aspects. The NMOC and NO_x concentrations in the HAOS data set (Table 8) are generally lower than in the HOMS data set (Table 9). The mean NMOC is 987.4 ppbC for Table 8 (with s = 545.0) and 2368.9 for Table 9 (with s = 1148.3). The difference between these two means is highly statistically significant. The same is true for NO_x, whose mean value is 89.5 ppb for Table 8 (with s = 56.0) and 139.2 for Table 9 (with s = 86.6). Another important difference is that the network of ozone monitors used in the HAOS was not the same as that used in the HOMS, although the two networks overlapped.

EVALUATION OF STANDARD EKMA

Comparison of Observed and Estimated Ozone

Figure 14 shows a scatterplot of OBS and EST for the HAOS data set. Region 1 contains two points, Region 2 has 11, and Region 3 has 48. (One of the points in Region 2 is plotted just below the line OBS = 0.8 EST.) Thus, about 3 percent of the cases are underpredicted, and the remainder satisfy the inequality OBS $\leq$ 1.2 EST. These percentages are similar to those previously obtained for the RAPS data (see Figure 4).

The observed and estimated ozone for the HOMS data set are compared in Figure 15. The figure shows that there are no points in Region 1, one point in Region 2, and the rest are in Region 3. The pattern of substantial overprediction depicted in Figure 15 is a consequence of the high values of NMOC and NO_x in the HOMS data set. The single point in Region 2 of Figure 15

TABLE 8. EVALUATION DATA SET FOR HOUSTON: HAOS DATA

Date (1977)	Precursors		Temperature (^{o}C)		Solar Radiation* 1000-1400 Average (ly/min)	Observed Maximum Ozone			EKMA Ozone Estimate (ppb)	
	NMOC (ppbC)	NO_x (ppb)	Daily Maximum	Morning Minimum		Station Name	Time (CDT)	O_3 (ppb)	Standard	City-Specific
1 May	330	57	27.1	17.8	0.68	Parkhurst	1500	79	137	58
2 May	470	53	29.4	20.0	0.49	Aldine	1400	90	159	65
3 May	1575	141	27.8	16.1	0.26	Parkhurst	1300	95	334	126
4 May	1543	89	28.3	20.6	0.65	Parkhurst	1200	62	277	114
6 May	1103	94	28.9	17.8	0.54	Aldine	1200	64	258	100
7 May	803	63	30.0	20.6	0.37	Parkhurst	1300	89	202	82
8 May	943	41	30.6	21.7	0.53	Parkhurst	1400	141	176	79
9 May	1457	188	27.8	19.4	0.31	Clinton	1700	103	354	117
12 May	625	46	30.0	15.0	0.67	Mae Drive	1300	169	167	71
19 May	1107	53	28.3	21.7	0.38	Parkhurst	1200	65	205	88
20 May	663	72	27.8	20.6	0.42	Parkhurst	1200	58	198	77
23 May	563	47	27.8	19.4	0.39	Fugua	1500	92	163	69
25 May	2396	203	31.1	16.1	0.64	Parkhurst	1200	302	435	65
27 May	1018	99	31.7	19.4	0.63	Fugua	1400	141	256	97
28 May	1650	69	31.7	17.2	0.68	Aldine	1500	80	252	108
29 May	1420	42	33.3	17.8	X	Aldine	1500	95	195	87
31 May	1545	110	33.9	17.8	0.79	Parkhurst	1500	264	302	120
1 Jun	957	89	31.1	20.6	0.41	Fugua	1500	280	242	94
2 Jun	1400	110	33.3	20.0	0.69	Aldine	1600	270	294	115
3 Jun	745	123	33.3	18.9	0.64	Aldine	1200	258	235	86
4 Jun	930	99	34.4	18.3	0.63	Lang	1200	256	248	92
5 Jun	990	100	35.0	18.9	0.68	Fugua	1400	180	254	95
6 Jun	460	54	35.0	21.1	0.69	Fugua	1300	186	159	64
8 Jun	555	43	32.2	16.7	0.71	Crawford	1500	309	157	64
9 Jun	1567	242	33.9	17.2	0.67	Crawford	1300	267	385	128
10 Jun	1087	91	33.9	18.3	0.57	Parkhurst	1500	119	254	99
12 Jun	700	43	29.4	22.8	0.24	Lang	1500	111	167	72
13 Jun	863	39	31.7	21.1	0.47	Lang	1300	77	169	75
16 Jun	1010	54	31.1	24.4	0.44	Westhollow	1700	62	202	86
17 Jun	833	45	32.8	22.8	0.66	Westhollow	1400	87	178	77
18 Jun	700	43	32.8	22.2	0.66	Fugua	1200	73	167	72
21 Jun	850	54	32.2	24.4	0.44	Westhollow	1800	70	193	81
23 Jun	1053	63	32.2	23.3	0.60	Westhollow	1600	108	218	91
24 Jun	1860	45	32.8	22.2	0.58	Fugua	1500	50	213	94
27 Jun	740	53	31.1	22.2	0.30	Westhollow	1500	94	185	77
28 Jun	625	58	33.9	21.7	0.62	Westhollow	1300	92	181	73

TABLE 8 (concluded)

Date (1977)	Precursors		Temperature (°C)		Solar Radiation* 1000-1400 Average (ly/min)	Observed Maximum Ozone			EKMA Ozone Estimate (ppb)	
	NMOC (ppbC)	NO_x (ppb)	Daily Maximum	Morning Minimum		Station Name	Time (CDT)	O_3 (ppb)	Standard	City-Specific
3 Jul	937	56	35.6	20.0	0.70	Aldine	1500	134	201	85
4 Jul	963	56	33.9	21.1	0.60	Jackrabbit	1500	149	203	86
5 Jul	1060	98	33.3	21.1	0.57	Parkhurst	1300	140	258	99
6 Jul	1165	115	32.2	22.2	0.53	Parkhurst	1200	143	281	106
7 Jul	1240	136	33.9	20.6	0.58	Parkhurst	1300	124	303	110
8 Jul	1150	134	33.3	21.1	0.45	Aldine	1200	112	293	105
9 Jul	660	101	34.4	21.7	0.62	Mae Drive	1300	152	215	79
10 Jul	473	60	33.9	22.2	0.61	Parkhurst	1300	63	165	66
12 Jul	677	98	35.6	23.3	0.67	Parkhurst	1200	107	216	79
15 Jul	873	70	34.4	21.1	0.55	MacGregor	1300	178	215	86
18 Jul	1597	82	32.8	21.1	0.48	Jackrabbit	1400	125	270	113
19 Jul	940	83	31.1	21.7	0.39	Parkhurst	1300	163	234	92
20 Jul	730	125	31.7	22.8	0.48	Parkhurst	1600	60	233	86
22 Jul	600	154	35.0	21.7	0.57	Mae Drive	1500	134	207	81
23 Jul	253	66	36.7	23.9	0.66	Mae Drive	1500	160	120	54
24 Jul	210	39	35.6	23.9	0.67	Aldine	1400	107	116	49
25 Jul	680	89	35.6	22.2	0.62	Aldine	1500	138	212	78
26 Jul	727	70	34.4	23.3	0.56	Clinton	1300	133	203	80
27 Jul	800	110	32.2	22.2	0.37	Mae Drive	1800	110	239	86
31 Jul	450	66	35.6	22.8	0.63	Parkhurst	1500	71	165	65
21 Oct	1206	126	26.7	18.3	0.37	Parkhurst	1600	93	293	108
23 Oct	570	58	26.7	15.6	0.40	MacGregor	1500	156	176	71
25 Oct	758	95	25.0	16.1	0.39	Clinton	1500	68	226	83
26 Oct	3710	378	28.9	13.9	0.53	Clinton	1400	122	523	223

*X denotes no data available.

TABLE 9. EVALUATION DATA SET FOR HOUSTON: HOMS DATA

Date (1978)	Precursors		Temperature (°C)		Observed Maximum Ozone			EKMA Ozone Estimate (ppb)	
	NMOC (ppbC)	NO_x (ppb)	Daily Maximum	Morning Minimum	Site	Time (CDT)	O_3 (ppb)	Standard	City-Specific
21 Sep	2208	88	31.1	22.8	Jackrabbit	1200	117	301	89
22 Sep	2489	61	31.1	23.3	Pearland	1500	119	260	79
25 Sep	2622	92	28.9	17.8	Pearland	1400	164	319	93
26 Sep	1411	80	29.4	20.0	Pearland	1200	182	259	81
27 Sep	867	88	28.9	21.7	Pearland	1400	106	233	75
29 Sep	1908	84	27.8	19.4	Pearland	1600	141	285	86
30 Sep	900	71	30.6	20.0	Darlinghurst	1500	105	218	71
1 Oct	1067	134	31.1	17.8	Hornwood	1400	310	285	87
2 Oct	2133	216	30.6	18.9	Aldine	1200	210	427	121
3 Oct	4542	122	28.3	21.7	Aldine	1200	100	408	117
4 Oct	2425	118	28.9	19.4	Aldine	1400	150	350	100
5 Oct	4900	157	30.0	21.1	Crawford	1300	190	467	133
7 Oct	1833	69	25.6	15.6	Darlinghurst	1600	163	258	80
9 Oct	2575	136	26.1	16.7	Aldine	1500	120	461	125
10 Oct	3133	194	27.8	15.0	Aldine	1200	150	461	125
11 Oct	3500	263	29.4	17.8	Sheldon	1300	133	537	143
12 Oct	1758	394	30.6	15.6	Hornwood	1500	120	448	117

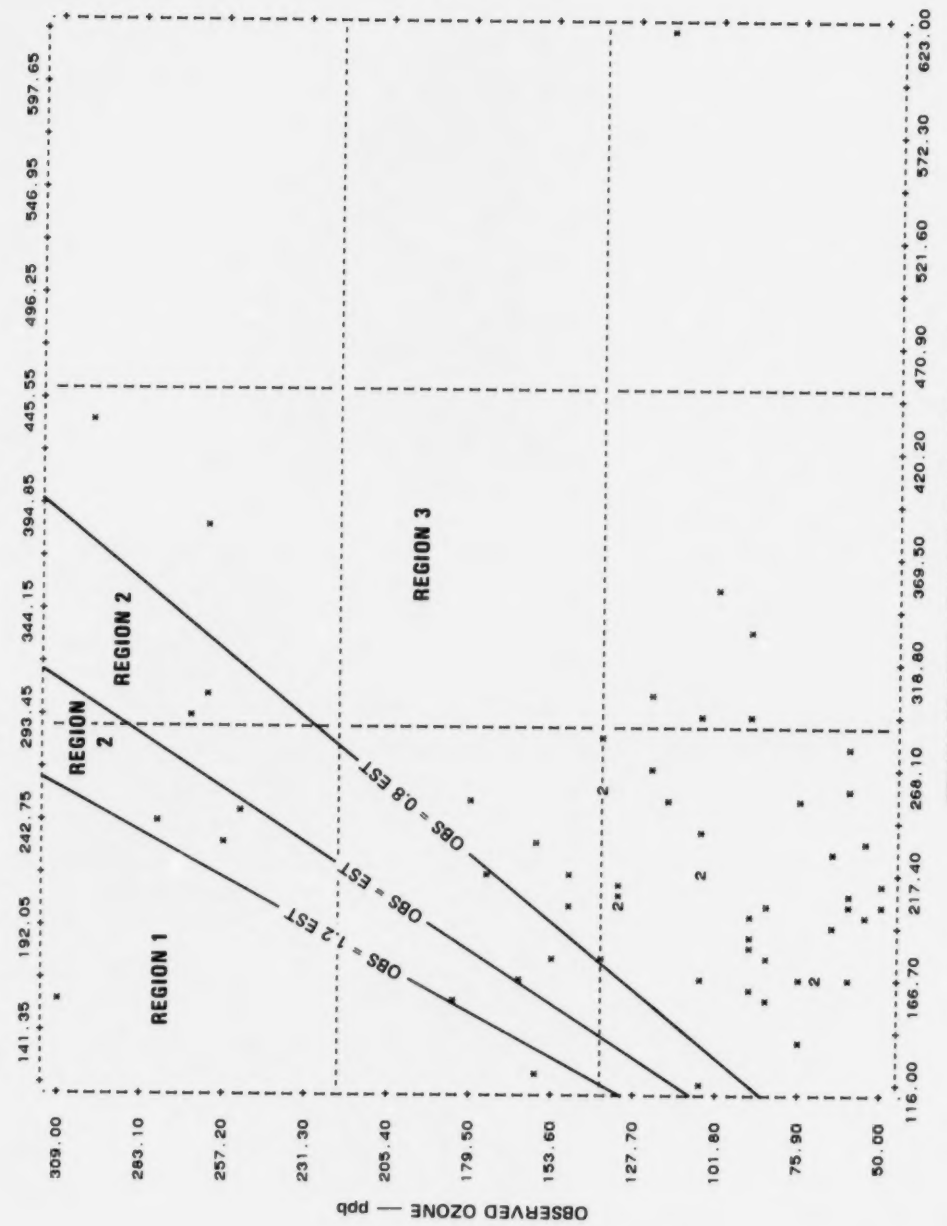


Figure 14. Scatterplot of observed ozone and standard-EKMA ozone estimate for HAOS data set. Number of points plotted is 61.

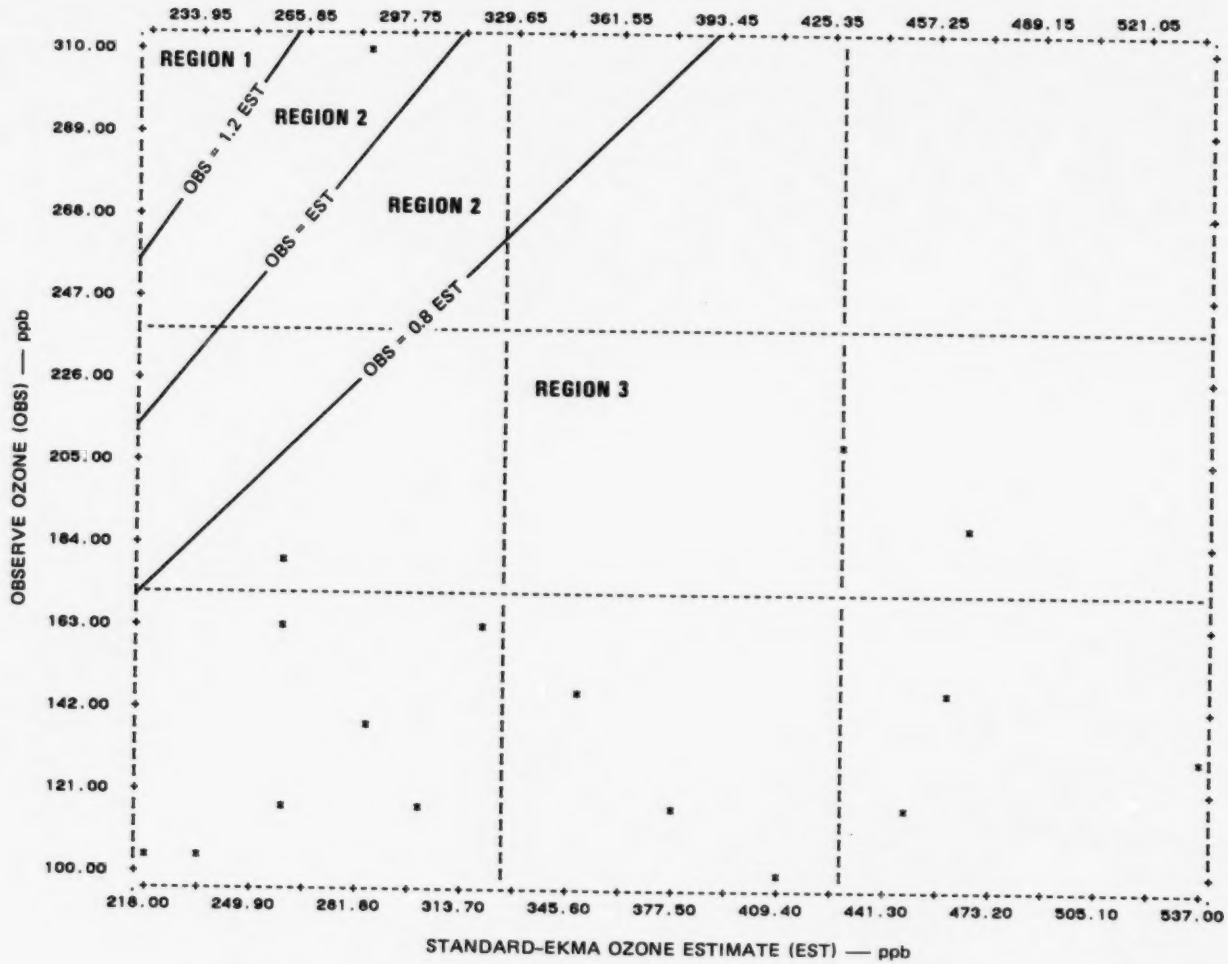


Figure 15. Scatterplot of observed ozone and standard-EKMA ozone estimate for HOMS data set. Number of points plotted is 17.

corresponds to 1 October, when the observed ozone was 310 ppb and the estimate was 285 ppb, hence R = 1.09. The ozone level of 310 ppb was the highest observed during the HOMS monitoring program (see Table 9). The NMOC/NO_x ratio for 1 October is about 8, the second lowest ratio in Table 9. (The lowest NMOC/NO_x ratio was 4.5, which occurred on 12 October.)

The two points in Region 1 of Figure 14 correspond to 8 June and 23 July (see Table 8). The OBS/EST ratio for these two days was 1.97 and 1.33, respectively. The ozone level of 309 ppb recorded on 8 June was the highest reported in the Houston area in 1977. It was measured at the Crawford monitoring site, which is located in downtown Houston. Analysis of the NMOC and NO_x data for 8 June showed that 0600-0900 measurements were not available for the Crawford site, and the precursor levels shown in Table 8 for this date were determined solely from data for Aldine and Parkhurst. The Crawford site tends to have the highest levels of NMOC and NO_x, which suggests the possibility that the precursor values used for 8 June may be low, thus inducing the underestimate. The data of 23 July exhibited the second lowest NMOC in Table 8. (The lowest was 210 ppbC, which occurred on 24 July). For 23 July, the NMOC was determined from measurements made at Aldine, Crawford, and Mae Drive. The 0600-0900 NMOC at Aldine and Crawford was 230 ppbC and at Parkhurst, 300 ppbC. Thus, although NMOC was quite low on this date, the data for the three sites appear to be consistent, and could well be valid.

The observed and estimated ozone shown in Figure 14 are correlated; the correlation is low ($r = 0.26$) but statistically significant ($p < 0.025$). The regression line is defined by OBS = 0.220 EST + 79.4, and has a standard error of 65.3, where all the units are ppb. By contrast, there is no statistically significant correlation between OBS and EST in Figure 15.

Relationships Between the Variables

Tables 10 and 11 show the correlation coefficients for selected variables for the HAOS and HOMS data sets, respectively. For convenience, the tables include variables associated with the city-specific ozone estimates; they will be dealt with in a later section.

Tables 10 and 11 show that temperature difference is correlated with both NO_x and OBS. OBS and NO_x are correlated in the HAOS data set (Table 10) but not in the HOMS data set (Table 11). The standard-EKMA ozone estimate, denoted by EST in the tables, is highly correlated with both NMOC and NO_x, the correlation being higher in the HAOS data set. Table 10 also shows that EST is slightly correlated with temperature difference in the HAOS data set but not in the HOMS data set (Table 11). Table 10 indicates that the ratio OBS/EST is correlated with 1/NMOC, NMOC, and temperature difference, which will be exploited in deriving a multiple regression equation for OBS/EST.

NMOC and NO_x are highly correlated in the HAOS data set, but there is no significant correlation between these two variables in the HOMS data set. For the HAOS data set, NO_x and NMOC are related by the regression equation $NO_x = 15.95 + 0.0745(NMOC)$. The significance level of the intercept is 0.065 and of the slope, 0.00001; the standard error of the regression is s = 38.96 ppb. Figure 16 shows the regression line and the evaluation region boundaries

TABLE 10. STATISTICAL PARAMETERS FOR SELECTED VARIABLES IN THE HAOS DATA SET*

Variable	Mean	Standard Deviation	Correlation Coefficient†								
			NO_x	1/NMOC	$1/NO_x$	Temp. Diff.	OBS	EST‡	OBS/EST	CSO3§	OBS/CSO3
NMOC (ppbC)	987.4	545.0	0.72	-0.70	0.46	0.23	--	0.90	-0.33	0.96	-0.30
NO_x (ppb)	89.5	56.0		-0.40	-0.80	0.31	0.27	0.92	--	0.85	--
1/NMOC ($ppbC^{-1}$)	0.0013	0.0008			0.41	--	--	-0.64	0.38	-0.68	0.33
$1/NO_x$ (ppb^{-1})	0.0141	0.0060				-0.27	-0.28	-0.75	--	-0.66	--
Temperature difference (°C)	11.6	2.7					0.61	0.28	0.44	0.26	0.49
OBS (ppb)	130.1	67.0						0.26	0.78	0.24	0.84
EST (ppb)‡	230.2	79.8							-0.28	0.98	--
OBS/EST	0.596	0.319								-0.29	0.99
CSO3 (ppb)§	90.3	26.6									-0.24
OBS/CSO3	1.49	0.770									1.0

*Statistics based on 61 points.

†Significant at 0.05 level or better. Dashes indicate that correlation is not significant.

‡Standard-EKMA estimate.

§City-specific EKMA estimate.

TABLE 11. STATISTICAL PARAMETERS FOR SELECTED VARIABLES IN THE HOMS DATA SET*

Variable	Mean	Standard Deviation	Correlation Coefficient†								
			NO_x	1/NMOC	$1/NO_x$	Temp. Diff.	OBS	EST‡	OBS/EST	CSO3§	OBS/CSO3
NMOC (ppbC)	2368.9	1148.3	--	-0.86	--	--	--	0.71	-0.54	0.75	-0.50
NO_x (ppb)	139.2	86.6		--	-0.87	0.71	--	0.77	--	0.72	--
1/NMOC ($ppbC^{-1}$)	0.0005	0.0003			--	--	--	-0.69	0.52	-0.69	0.46
$1/NO_x$ (ppb^{-1})	0.0092	0.0040				-0.59	--	-0.89	--	-0.86	--
Temperature difference (°C)	10.1	2.3					0.44	--	--	--	--
OBS (ppb)	151.8	51.7						--	0.81	--	0.86
EST (ppb)‡	346.7	96.3							-0.56	0.99	-0.48
OBS/EST	0.47	0.20								-0.53	0.99
CSO3 (ppb)§	100.2	22.1									-0.46
OBS/CSO3	1.57	0.63									1.0

*Statistics based on 17 points.

†Significant at 0.05 level or better. Dashes indicate that correlation is not significant.

‡Standard-EKMA estimate.

§City-specific EKMA estimate.

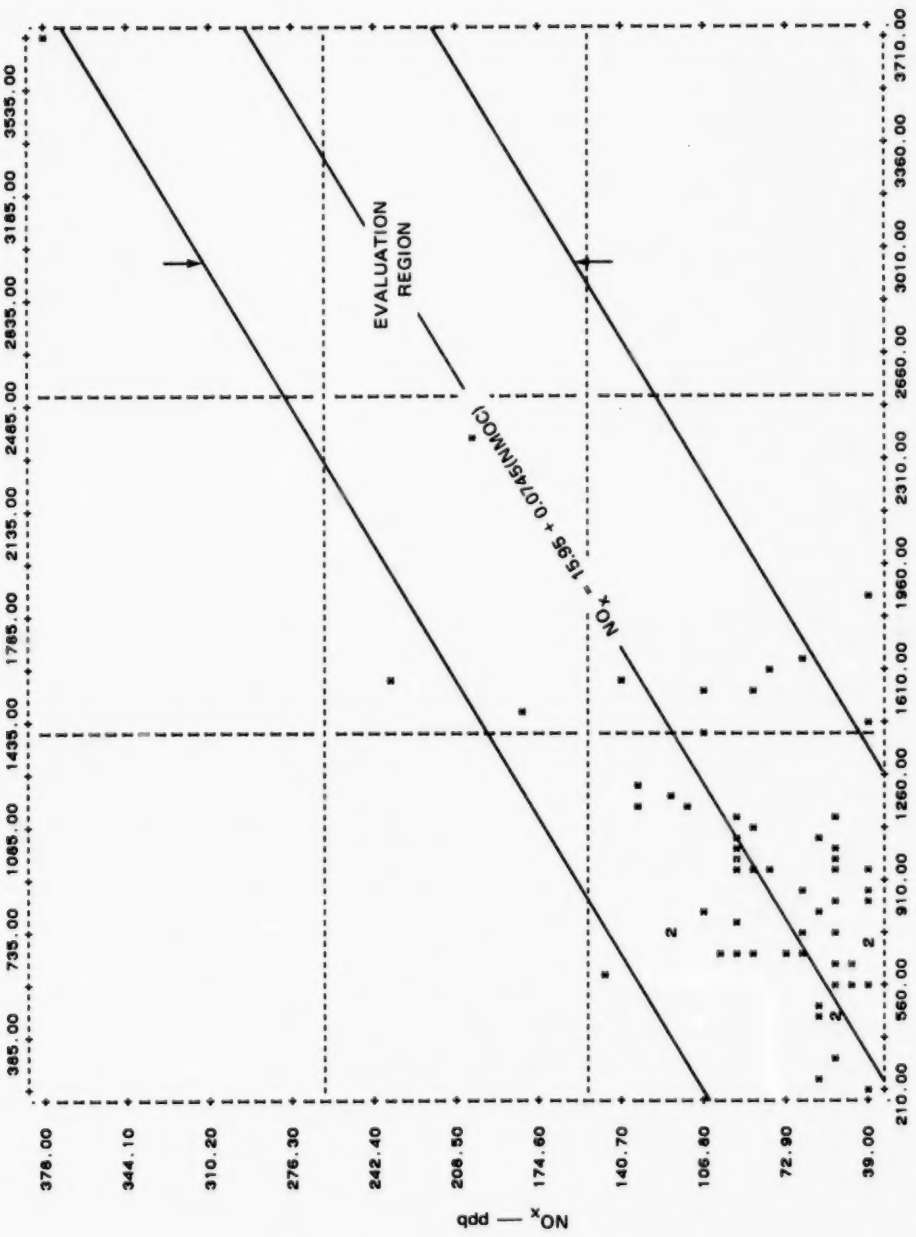


Figure 16. Scatterplot of NO_x and NMOC for the HAOS data set. Number of points plotted is 61.

for the HAOS data. The lack of correlation between NMOC and NO_x in the HOMS data set is evident in the scatterplot of Figure 17.

Definition of Accuracy Regions

A multiple regression equation was derived for OBS/EST as a function of NMOC and temperature difference (denoted by DT) for the HAOS data set. The equation is defined by

$$\text{OBS/EST} = 0.1145 - 0.0002681\ (\text{NMOC}) + 0.06425\ (\text{DT}) \qquad (6)$$

where the units of the variables were defined in Table 10. The multiple correlation coefficient is $r = 0.63$ and the standard error of the regression is $s = 0.25$. The significance level of the regression and of the coefficients of NMOC and DT is $p < 0.0001$, but the constant term is not statistically significant at the 0.05 level. An expression analogous to Eq. 6 was not obtained for the HOMS data set because the sample size is too small to yield meaningful results.

Following the procedures previously described in Sections Two and Three, a plot of the accuracy probability was obtained using Eq. 6 after setting DT = 17°C, which is its mean value plus two standard deviations. Figure 18 displays the probability curves for the three accuracy regions as a function of the variable $Z = -0.0002681\ (\text{NMOC})$. The figure shows that $P(R \leq 1.2) > 0.5$ for $Z < -0.175$, which corresponds to NMOC > 654 ppbC. This reflects the standard EKMA's tendency to overpredict. The propensity for overprediction becomes even more pronounced if the value of DT in Eq. 6 is lowered from 17°C to its mean value of 11.6°C. To illustrate, consider the circled point in Figure 18, which is located at $Z = -0.45$. Setting DT = to 11.6°C would shift the curves in Figure 18 to the right, changing the Z-coordinate of the circled point from -0.45 to -0.10. Thus, nearly all of Figure 18 to the right of the circled point would be truncated, and $P(R < 0.8) > 0.5$ over the entire range of NMOC and NO_x concentrations.

Figure 18 shows that the curve for $P(0.8 \leq R \leq 1.2)$ is flattened and spread out, in sharp contrast to the relatively narrow curve shown for St. Louis in Figure 7. (Note, however, that the Z-scales are not equal in Figures 7 and 18.) The maximum value of $P(0.8 \leq R \leq 1.2)$ is 0.58 in Figure 18, compared to the maximum of 0.83 in Figure 7. The curves in Figures 7 and 18 differ in shape because the standard error of the regression is smaller for Eq. 2 ($s = 0.15$) than for Eq. 6 ($s = 0.25$). The maximum probability of 0.58 in Figure 18 occurs in the neighborhood of $Z = -0.21$, which corresponds to NMOC = 783 ppbC. The shaded region in Figure 18 is where $P(0.8 \leq R \leq 1.2) \geq 0.50$ and is defined by $-0.35 < Z < -0.060$, which corresponds to $224 < \text{NMOC} < 1305$ ppbC.

Figure 19 displays constant-Z lines on the NMOC-NO_x plane, along with the evaluation region previously defined in Figure 16. The shaded area corresponds to the region where $P(0.8 \leq R \leq 1.2) \geq 0.50$. Inside the evaluation region, (NMOC, NO_x) combinations that fall within and to the right of the shaded area satisfy the inequality $P(R \leq 1.2) > 0.60$. Points to the left of the shaded area satisfy the relation $P(R > 1.2) > 0.42$, and thus are more

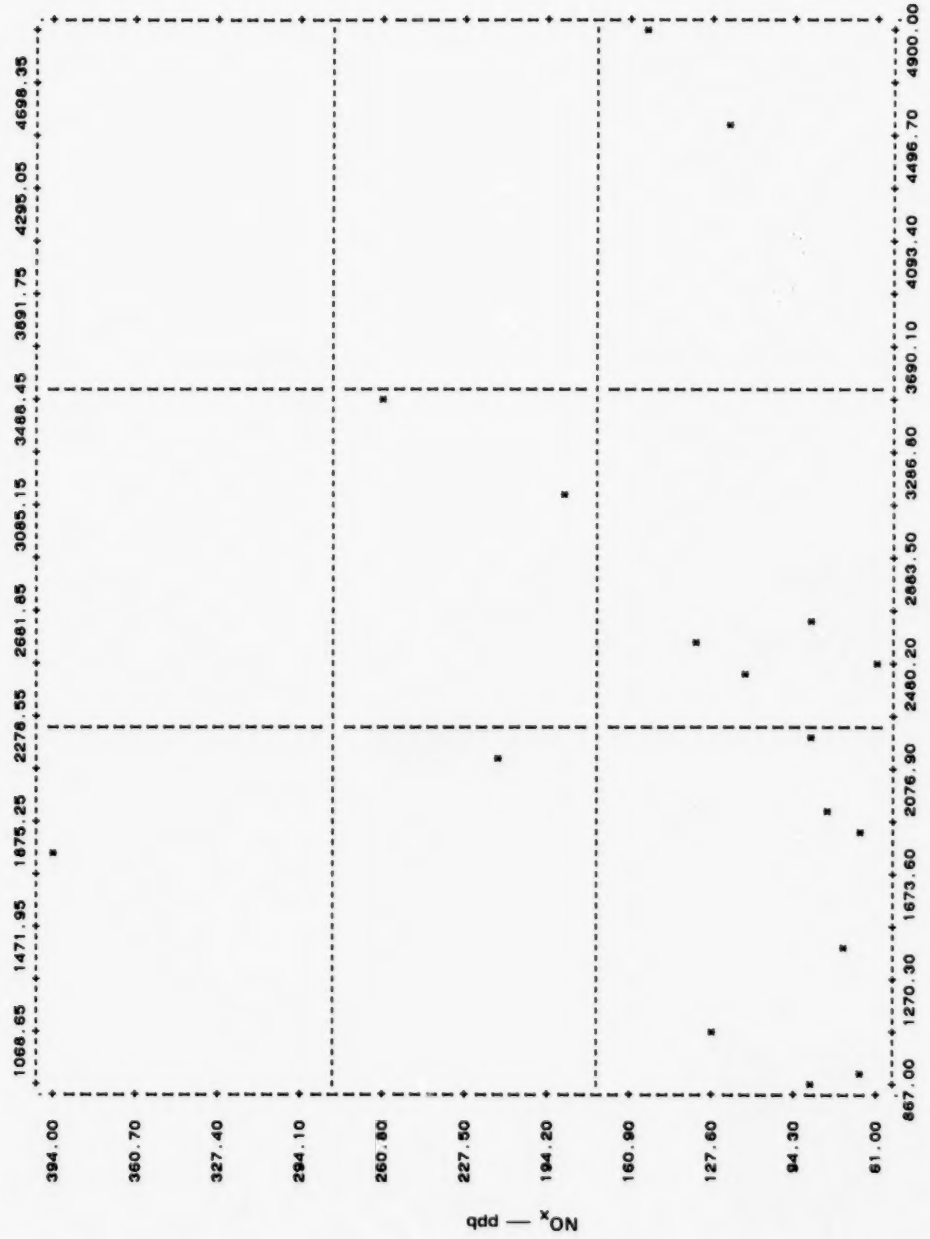


Figure 17. Scatterplot of NO_X and NMOC for the HOMS data set.
Number of points plotted is 17.

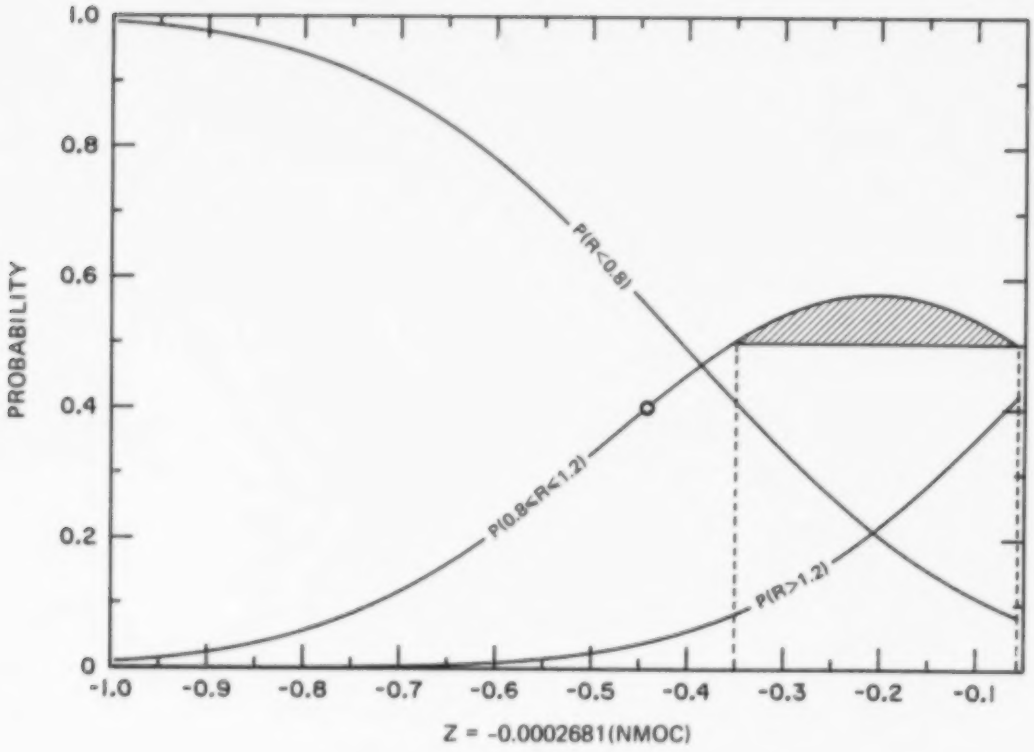


Figure 18. Accuracy probability for standard-EKMA ozone estimates for the HAOS data set. Temperature difference set to 17.0° C.

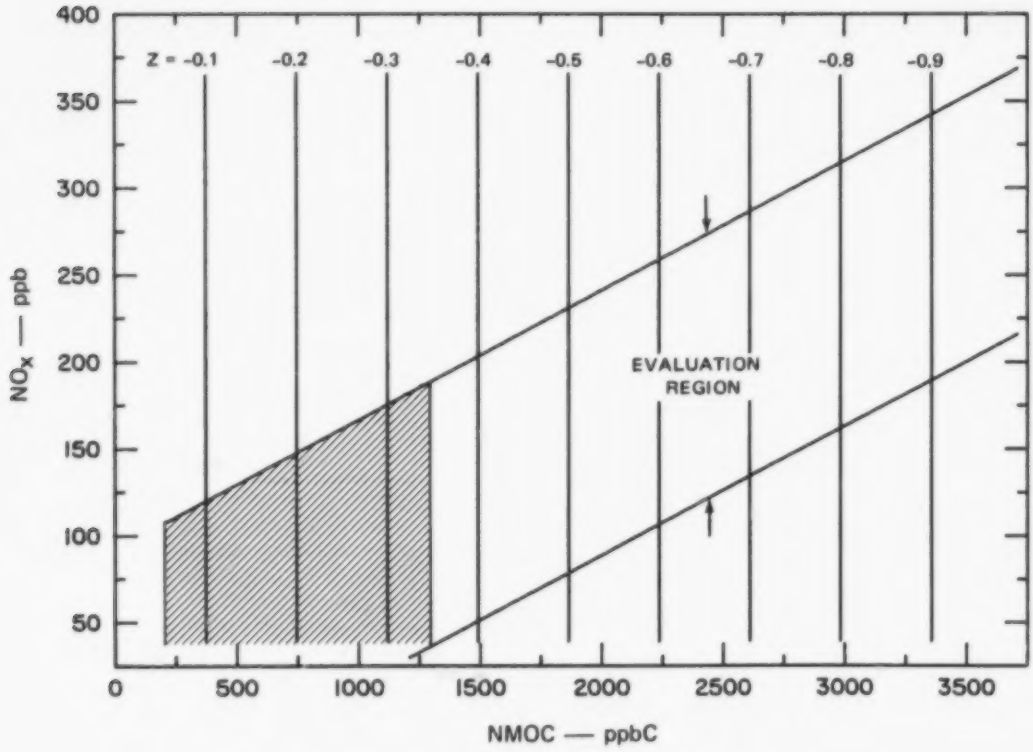


Figure 19. Constant-Z plot for standard-EKMA ozone estimates for the HAOS data set. Shading indicates area where ozone estimates tend to be most accurate, assuming a temperature difference of 17.0° C.

likely to yield underestimates of observed ozone. The tendency of the standard EKMA to overpredict ozone is reflected in the large difference in the respective sizes of the regions that flank the shaded area in Figure 19. If the figure were redrawn to correspond to a value of DT = 11.6°C, the shaded area of Figure 19 would be confined to a narrow slice at the lower left hand corner of the diagram, and the region of overprediction would cover essentially the entire evaluation region.

EVALUATION OF CITY-SPECIFIC EKMA

City-Specific Ozone Estimates

Table 12 shows the input parameters used to obtain the city-specific ozone estimates for both the HAOS and HOMS data sets. The criteria for defining this input follow those used with the St. Louis data. Inversion height data were derived from soundings obtained for 8 June 1977. The dates shown in Table 12 correspond to the days with the highest observed ozone for each data set.

As in the St. Louis case, the C-S and standard-EKMA ozone estimates were highly correlated for both HAOS and HOMS. Tables 10 and 11 show that the correlation coefficient between C-S and standard-EKMA estimates is r = 0.98 for the HAOS, and r = 0.99 for the HOMS. The regression line for the HAOS data is CSO3 = 15.02 + 0.33 * EST, and that for the HOMS is CSO3 = 21.02 + 0.23 * EST, where the units are ppb.

Comparison of Observed Ozone and City-Specific Estimates

Figures 20 and 21 show scatterplots of observed and estimated ozone for the HAOS and HOMS data, respectively. Whereas the standard-EKMA estimates showed a marked tendency to overpredict, the C-S estimates show the opposite: The majority of the points in both figures are in Region 1, the region of underprediction. For the HAOS, Figure 20 has 33 points (54 percent) in Region 1, 16 points (26 percent) in Region 2, and 12 points (20 percent) in Region 3. Figure 21 for the HOMS has 12 points (71 percent) in Region 1, 5 points (29 percent) in Region 2, and none in Region 3. Observed and estimated ozone are correlated for the HAOS data; the correlation is low (r = 0.24) but statistically significant ($p < 0.04$). No significant correlation exists between the two quantities in the HOMS data.

Definition of Accuracy Regions

A multiple regression equation analogous to Eq. 6 was derived for the ratio OBS/CSO3 for the HAOS data set. The equation is

$$\text{OBS/CSO3} = 0.1718 - 0.0006191\ (\text{NMOC}) + 0.1664\ (\text{DT}) \quad . \qquad (7)$$

The units and symbols in Eq. 7 were previously defined. The multiple correlation coefficient is r = 0.65 and the standard error is s = 0.60. As in Eq. 6, the significance level of the coefficients of NMOC and DT is $p < 0.0001$, but

TABLE 12. SUMMARY OF INPUT PARAMETERS FOR OBTAINING CITY-SPECIFIC EKMA OZONE ESTIMATES FOR HOUSTON

Parameter	Value
Date	
HAOS	8 June 1977
HOMS	1 October 1978
Location	Latitude 29.75°N, Longitude 95.40°W
Inversion height data	
Initial height (m)	105
Final height (m)	1510
Starting time of rise	0700
Ending time of rise	1400
Post-0800 emissions by ending hour	
0900	
NMOC	0.23
NO_x	0.23
1000	
NMOC	0.07
NO_x	0.05
1100	
NMOC	0.08
NO_x	0.06
1200	
NMOC	0.03
NO_x	0.03
1300	
NMOC	0.04
NO_x	0.03
1400	
NMOC	0.01
NO_x	0.03
Reactivity, NO_2/NO_x	0.33
Background ozone concentration (ppm)	0.040

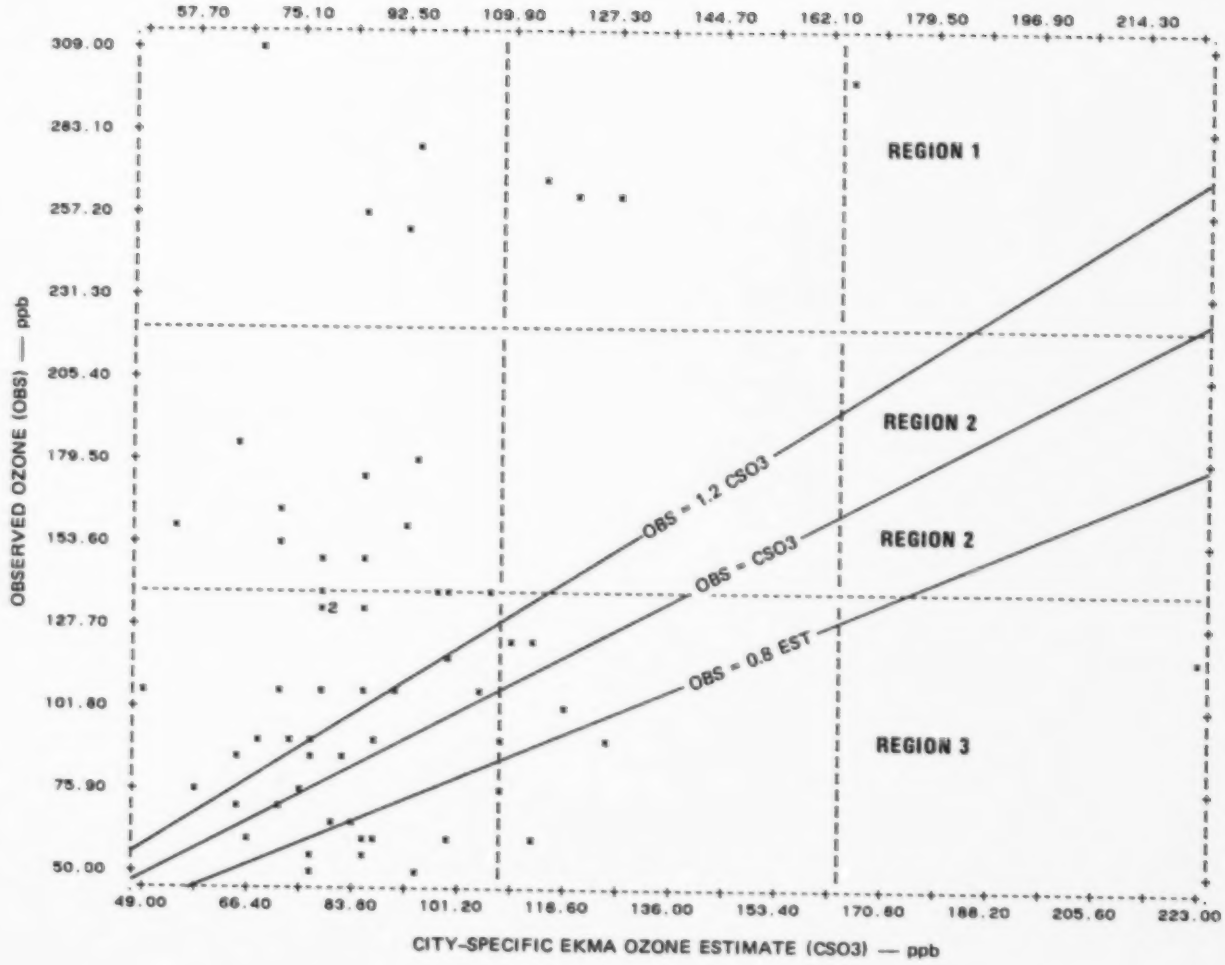


Figure 20. Scatterplot of observed ozone and city-specific EKMA ozone estimates for the HAOS data.

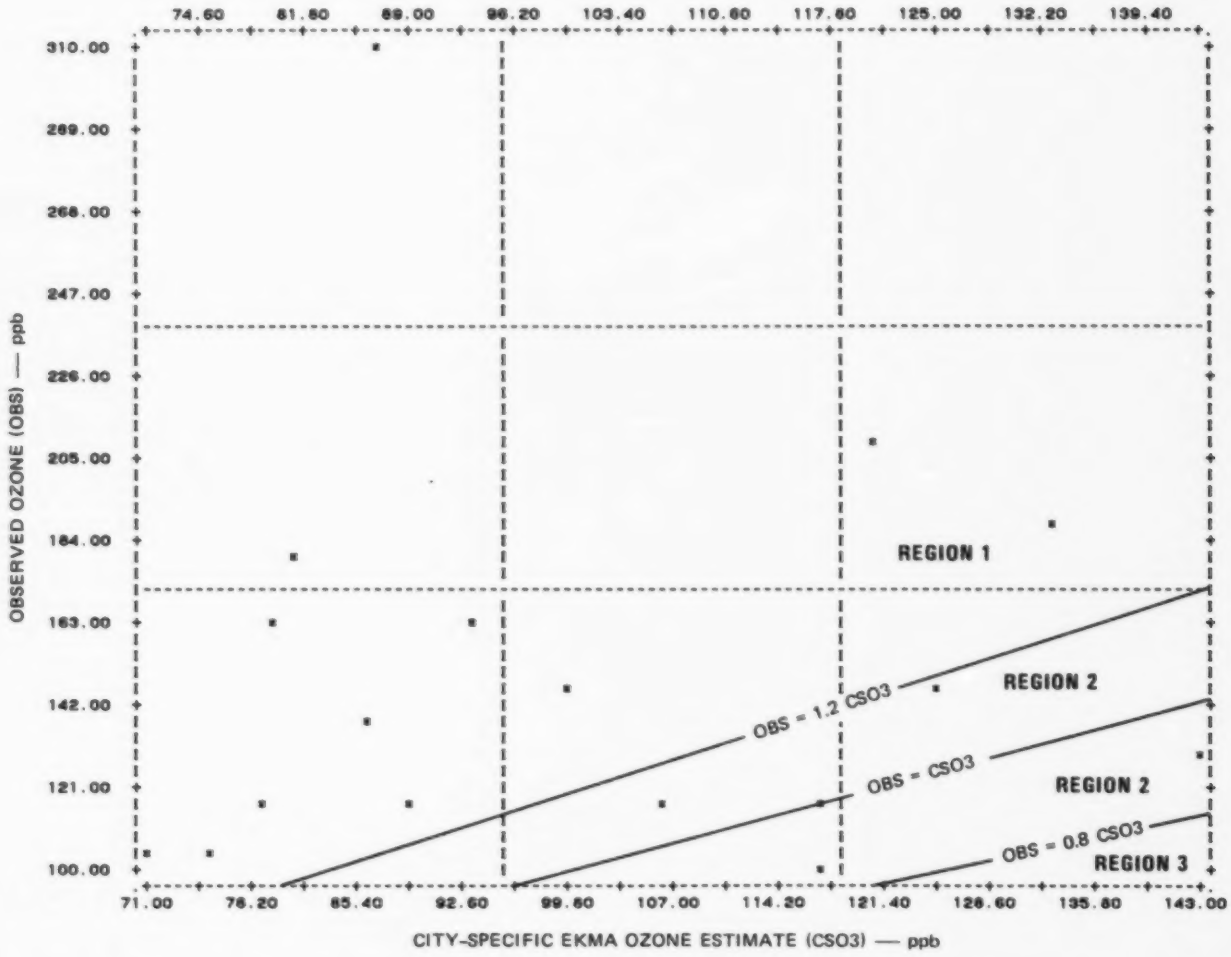


Figure 21. Scatterplot of observed ozone and city-specific EKMA ozone estimates for the HOMS data.

the constant term is not statistically significant at the 0.05 level. Equations 6 and 7 have the same form, but different coefficients; both equations explain approximately the same amount of variance (about 40 percent). However, the standard error of Eq. 7 is more than twice as big as that for Eq. 6. Thus, there is more scatter, and less precision, associated with Eq. 7. As a result, the accuracy probability plot will show distributions that are flat and broad.

Figure 22 displays the probability curves for the three accuracy regions as a function of the variable Z = -0.0006191 (NMOC). The curves are computed for DT = 11.6°C, which is its mean value. The probability $P(R \leq 1.2)$ exceeds 0.5 for $Z < -0.90$, which corresponds to NMOC > 1454 ppbC. Thus, the C-S ozone estimates can be considered to be upper bounds generally for high values of NMOC. In contrast to the corresponding plots for St. Louis (cf. Figure 12), the probability curve for Region 2 is very flat and broad. This is a consequence of the large standard error associated with Eq. 7. The curve for $P(0.8 \leq R \leq 1.2)$ covers the entire range of NMOC concentrations, but the probability is low that a given NMOC value will fall in Region 2. The highest probability of an accurate prediction is about 0.26, and occurs for Z = -1.11, which corresponds approximately to NMOC = 1793 ppbC.

Figure 23 depicts constant-Z lines on the NMOC-NO_x plane, along with the evaluation region. The area to the left of the line Z = -1.0 is associated with underprediction and that to the right of the line Z = -1.25 with overprediction. The area between the lines Z = -1.0 and Z = -1.25 is associated with the highest probability of an accurate prediction, but, as shown in Figure 22, the probability is low.

DISCUSSION

For both HAOS and HOMS data, the EKMA substantially overpredicted in the standard mode, and underpredicted in the city-specific mode. As a result, the probability of an accurate prediction for the HAOS data was generally low. In the city-specific case, the multiple regression fit to the ratio OBS/CSO3 had a large standard error that is indicative of a low-precision fit. In general, it appears that in either mode the EKMA tends to be a low-accuracy predictor of ozone for the Houston area. If it is desired to obtain an upper bound for the maximum potential ozone in the Houston area, then the standard-EKMA mode is the appropriate choice, because it has a low probability of underestimating the ozone level.

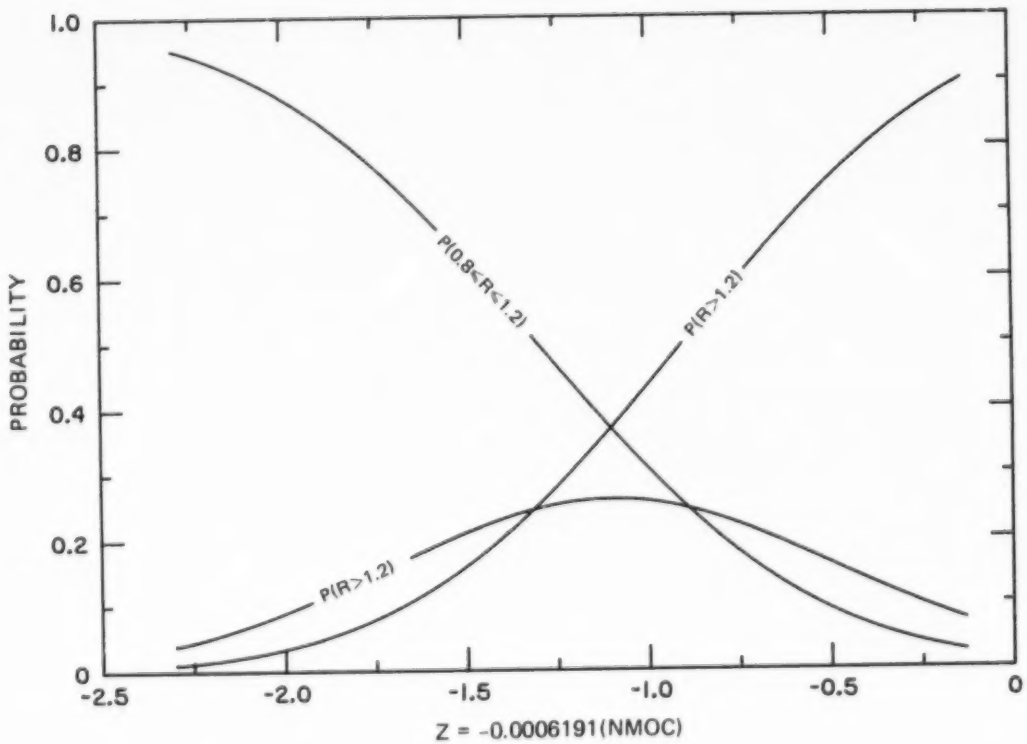


Figure 22. Accuracy probability for city-specific EKMA ozone estimates for the HAOS data set. Temperature difference is set to its mean value of 11.6°C.

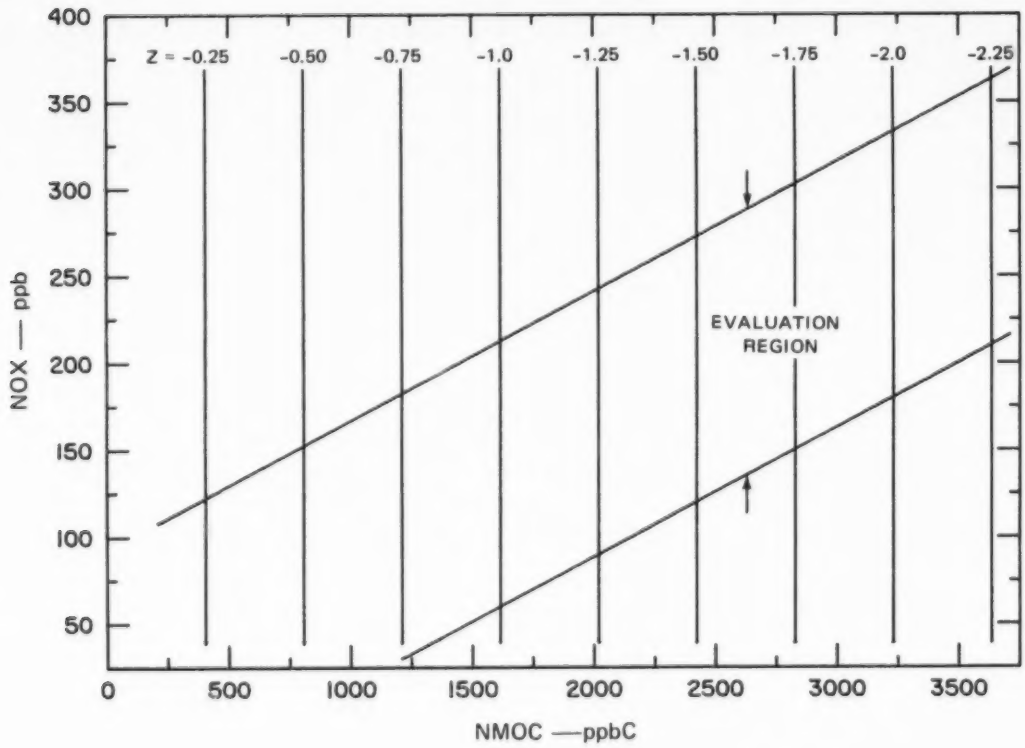


Figure 23. Constant-Z plot for city-specific EKMA ozone estimates for the HAOS data set. Temperature difference is set to its mean value of 11.6°C.

SECTION 5

EKMA EVALUATION FOR THE PHILADELPHIA AREA

DATA REVIEW AND ANALYSIS

The EKMA has also been evaluated using data collected in the Philadelphia area during the period July-September 1979. Figure 24 shows the 17 air-quality monitoring stations used during the field program.

The statistical analyses performed using the Philadelphia data found relationships similar to those found for St. Louis and Houston. The daily maximum ozone was found to have a significant positive correlation with the daily maximum temperature and the morning minimum to afternoon-maximum temperature difference. A low negative correlation was found between the daily peak ozone concentration and the 0600-1400 average wind speed.

The 0600-0900 EDT spatial average NMOC and NO_x concentrations were calculated using data for three source-region monitors located in downtown Philadelphia at South Broad and Spruce Streets (Site 13), at the Franklin Institute (Site 14), and at the American Meteorological Society Laboratory (Site 15).

DEFINITION OF EVALUATION DATA SET

Based on the data analysis, the following criteria were used to select the evaluation data set for Philadelphia:

- Maximum daily temperature at least 24°C.
- Difference between maximum and morning minimum temperature at least 7°C.
- 0600-0900 areawide average NO_x at least 34 ppb.
- 1000-1400 average wind speed less than 5 m/s.

Note that the first three criteria are similar to those used in connection with the Houston data. (Recall again that these meteorological conditions are always present whenever the daily maximum ozone is at least 100 ppb.)

Application of these criteria yielded an evaluation data set for Philadelphia that contains 29 days. Table 13 shows the data set including date, precursor levels, temperature, and observed and estimated ozone. The observed maximum O_3 exceeded 120 ppb on 12 of the 29 days.

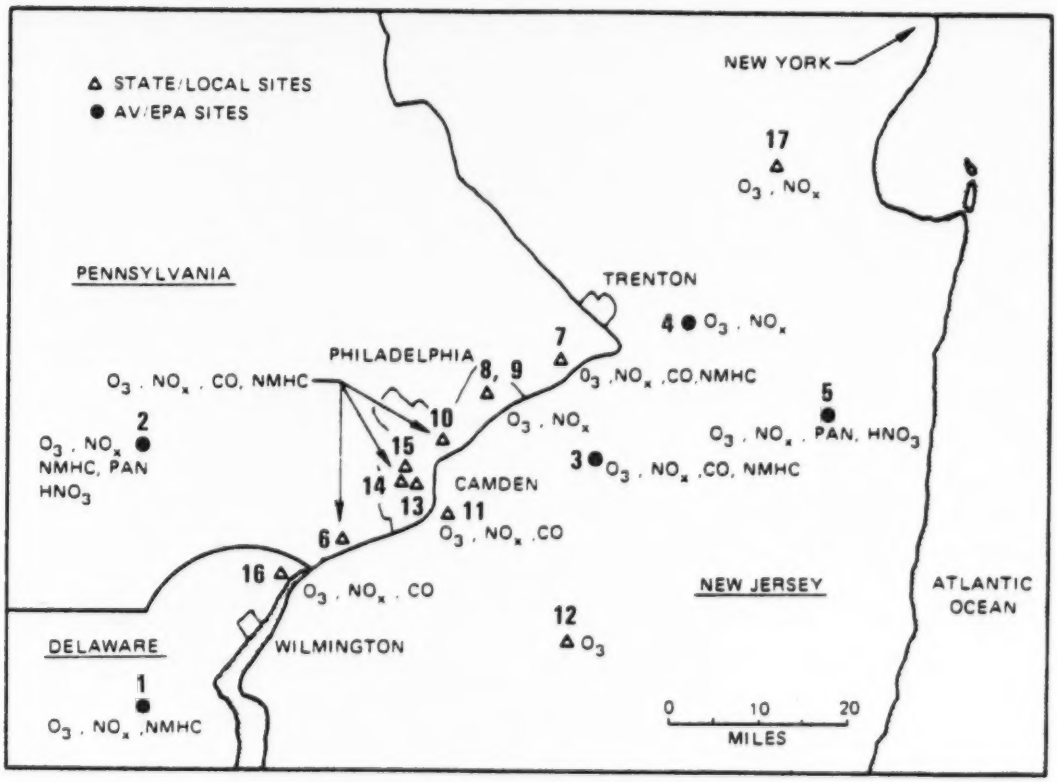


Figure 24. Air quality monitoring network for the Philadelphia area.

TABLE 13. EVALUATION DATA SET FOR PHILADELPHIA

Date (1979)	Precursors		Temperature (°C)		Observed Maximum Ozone			EKMA Ozone Estimate (ppb)	
	NMOC (ppbC)	NO_x (ppb)	Daily Maximum	Morning Minimum	Site*	Time (EDT)	O_3 (ppb)	Standard	City-Specific
7 Jul	800	97	24	11	11	1400	105	232	230
8 Jul	1267	133	26	12	10	1600	126	304	285
9 Jul	1250	90	26	13	10	1700	109	263	239
10 Jul	1567	93	25	14	13	1600	90	284	243
11 Jul	283	51	26	17	11	1700	110	124	137
12 Jul	783	81	30	18	11	1400	147	219	215
13 Jul	1083	144	31	19	6	1300	183	292	282
14 Jul	183	43	28	20	2	1200	84	101	110
17 Jul	300	53	30	20	5	1500	161	129	141
19 Jul	308	43	26	18	15	1700	160	123	137
20 Jul	450	87	27	17	2	1600	101	188	181
22 Jul	889	63	26	16	5	1600	141	208	199
23 Jul	389	48	26	19	2	1300	76	144	152
24 Jul	983	75	27	18	7	1500	120	229	216
25 Jul	767	57	29	20	5	1500	77	192	188
28 Jul	217	52	28	19	3	1300	146	11	120
31 Jul	317	52	29	20	2	1400	147	133	145
1 Aug	783	81	32	22	7	1500	114	219	215
10 Aug	783	63	31	19	15	1600	170	200	195
13 Aug	267	38	23	13	2	1700	70	116	127
14 Aug	983	96	27	17	5	1500	115	250	239
21 Aug	200	69	26	17	7	1700	99	80	104
22 Aug	900	111	28	16	14	1500	140	253	247
23 Aug	300	88	28	16	7	1600	77	121	135
25 Aug	225	36	28	21	7	1600	116	102	118
27 Aug	383	53	28	21	7	1600	101	146	155
28 Aug	422	59	27	20	7	1600	118	156	165
30 Aug	556	80	29	21	3	1400	126	189	195
31 Aug	367	91	27	19	11	1600	124	153	158

*Site numbers are keyed to Figure 24.

EVALUATION OF STANDARD EKMA

Comparison of Observed and Estimated Ozone

Figure 25 shows a scatterplot of OBS as a function of EST, which is the mnemonic for the standard-EKMA ozone estimate. Of the 29 points plotted, four (14 percent) are in Region 1, five (17 percent) are in Region 2, and the remainder in Region 3. Note, however, that the four points in Region 1 are not grossly underpredicted.

Thus, as in St. Louis and Houston, the standard-EKMA shows a marked tendency to overpredict. Unlike the St. Louis case, there is no statistically significant correlation between OBS and EST.

Relationships Between the Variables

Table 14 shows the pairwise correlation coefficients, mean, and standard deviation for selected variables in the Philadelphia data set; the table contains variables for both standard- and city-specific EKMA ozone estimates. The table shows that EST is highly correlated with both NMOC and NO_x, which was expected. However, there is no significant correlation between OBS and NMOC and OBS and NO_x. Maximum temperature is correlated only with OBS. The variable 1/NMOC correlates with EST and with OBS/EST because NMOC does.

Figure 26 shows a scatterplot of NMOC and NO_x. The plot clearly portrays the linear correlation between these two variables. As Table 14 indicates, the correlation coefficient is $r = 0.70$, and is highly statistically significant ($p < 0.00001$). The figure also shows the regression line that relates NO_x and NMOC. The regression equation is $NO_x = 42.26 + 0.0501(NMOC)$, and has a standard error $s = 19.8$ ppb. As in previous cases, the regression equation has been used to define an evaluation region for the NMOC-NO_x plane that is shown in Figure 26.

Definition of Accuracy Regions

The multiple regression equation derived for the OBS/EST ratio is

$$OBS/EST = -0.653 + 162.45/NMOC + 0.0365(T) \quad , \qquad (8)$$

where T denotes maximum temperature in ^{o}C, and the other symbols and units are as previously defined. The multiple regression coefficient for Eq. 8 is $r = 0.79$ and the standard error is $s = 0.198$. The coefficients of NMOC and T are statistically significant at the 0.05 level, but the constant is not. Neither NO_x nor $1/NO_x$ appear in Eq. 8 because both are correlated with 1/NMOC and do not provide any statistically significant improvement in the amount of variance explained.

Figure 27 displays the accuracy probability plot for OBS/EST as a function of $Z = 162.45/NMOC$ for the mean value of T, which is $27.5^{o}C$. The probability $P(R \leq 1.2)$ exceeds 0.5 for $Z < 0.85$, which corresponds to NMOC > 191 ppbC. The shaded area in Figure 27 defines the region with the highest probability of an accurate prediction, i.e. where $0.8 \leq OBS/EST \leq 1.2$. In the

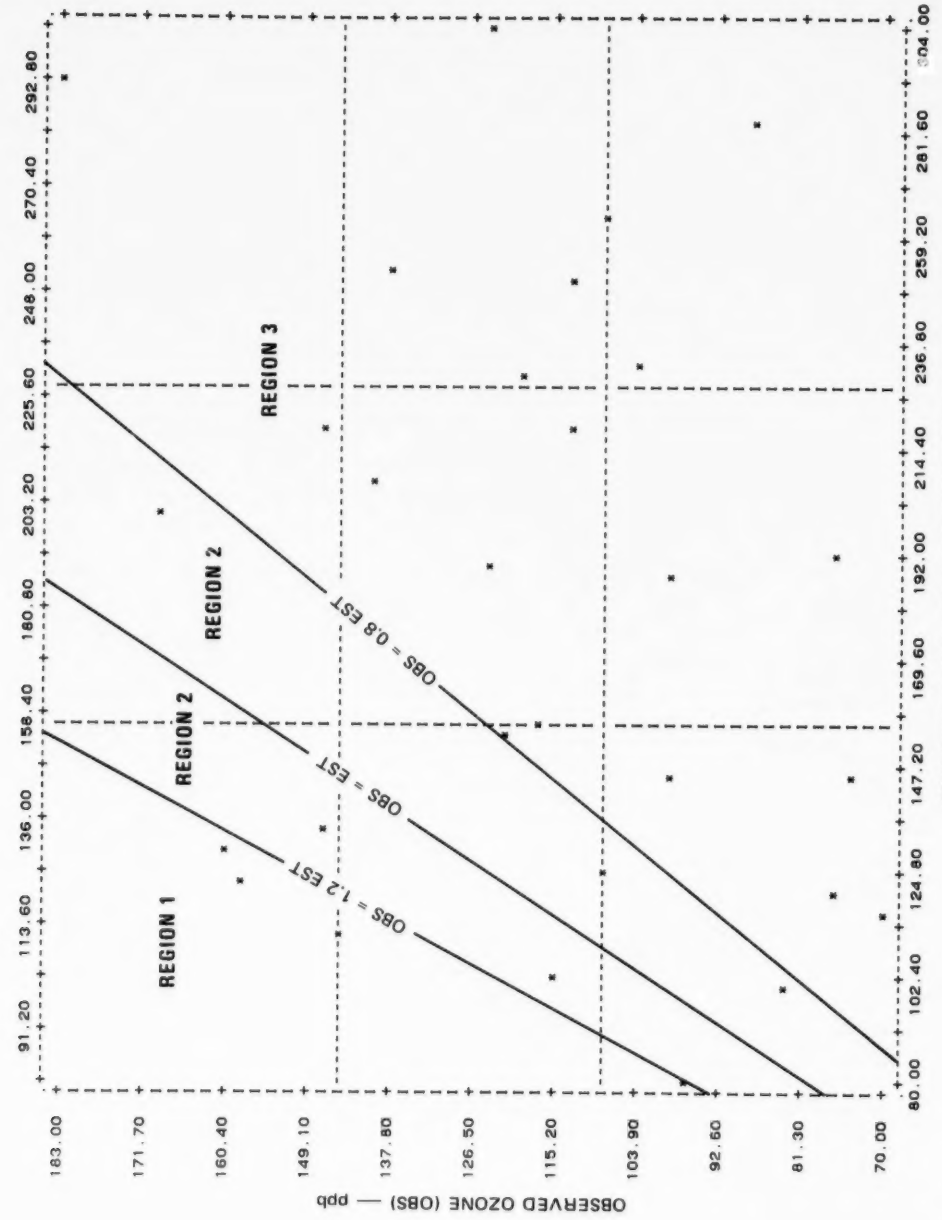


Figure 25. Scatterplot of observed ozone and standard-EKMA ozone estimate for Philadelphia.

TABLE 14. MEANS, STANDARD DEVIATIONS, AND CORRELATIONS FOR SELECTED VARIABLES IN THE PHILADELPHIA DATA SET

Variable	Mean	Standard Deviation	Correlation Coefficient*							
			NO_x	1/NMOC	Maximum Temperature	OBS	EST	OBS/EST	CSO3	OBS/CSO3
NMOC (ppbC)	620.9	379.8	0.70	-0.88	--	--	0.96	-0.71	0.92	-0.65
NO_x (ppb)	73.3	27.2		-0.65	--	--	0.81	-0.54	0.84	-0.51
1/NMOC ($ppbC^{-1}$)	0.0023	0.0014			--	--	-0.92	0.75	-0.92	0.66
Temperature difference (^oC)	27.5	2.1				0.52	--	--	--	--
OBS (ppb)	119.1	29.8					--	0.42	--	0.54
EST (ppb)†	181.4	64.1						-0.74	0.99	-0.66
OBS/EST	0.73	0.29							-0.71	0.98
CSO3 (ppb)‡	181.8	51.9								-0.64
OBS/CSO3	0.70	0.24								1.0

*All correlations significant at the 0.05 level or better.
Dashes denote that correlation is not statistically significant.

†Standard-EKMA ozone estimate.

‡City-specific EKMA ozone estimate.

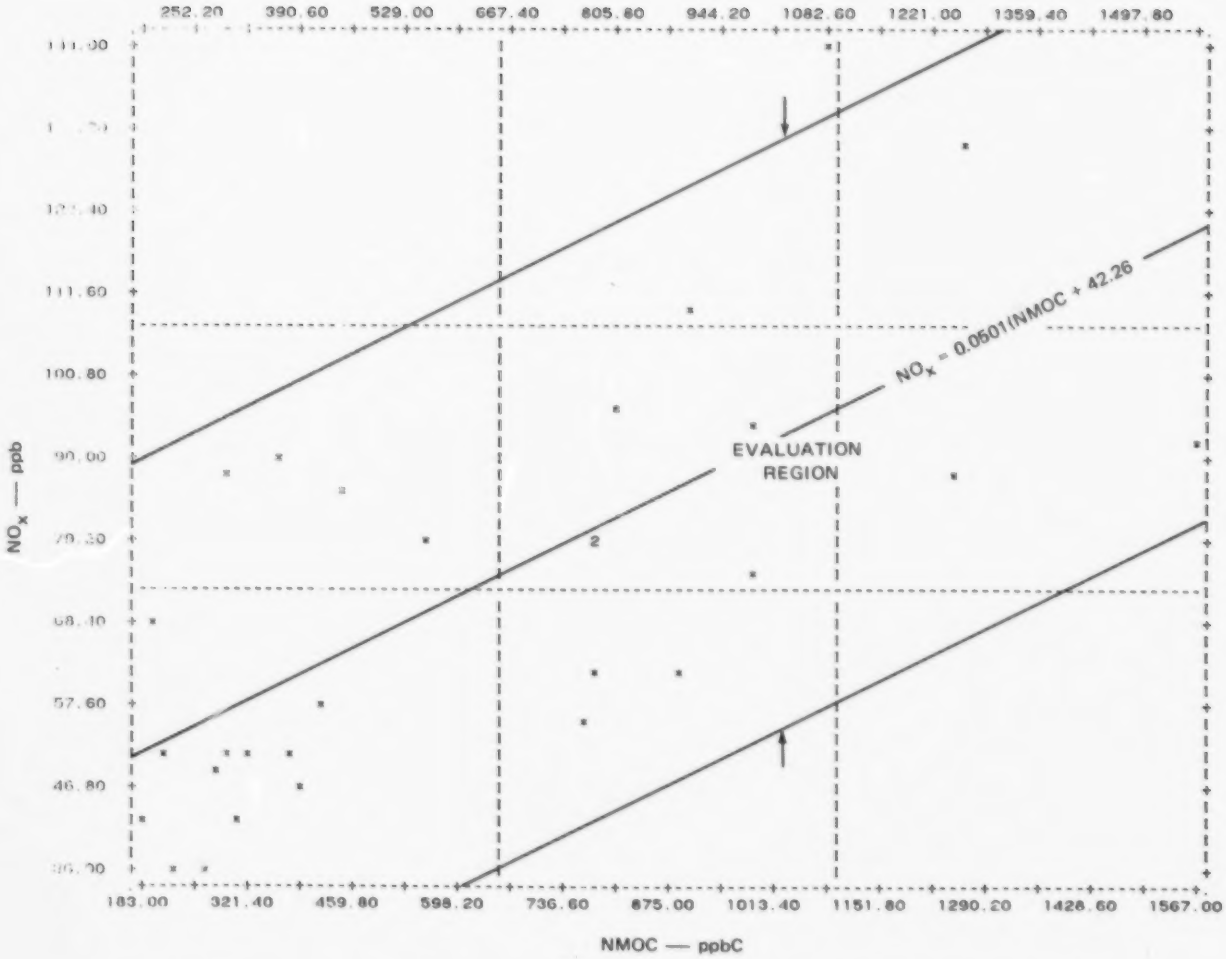


Figure 26. Scatterplot of 0600–0900 (EDT) concentrations of NMOC and NO_x for Philadelphia.

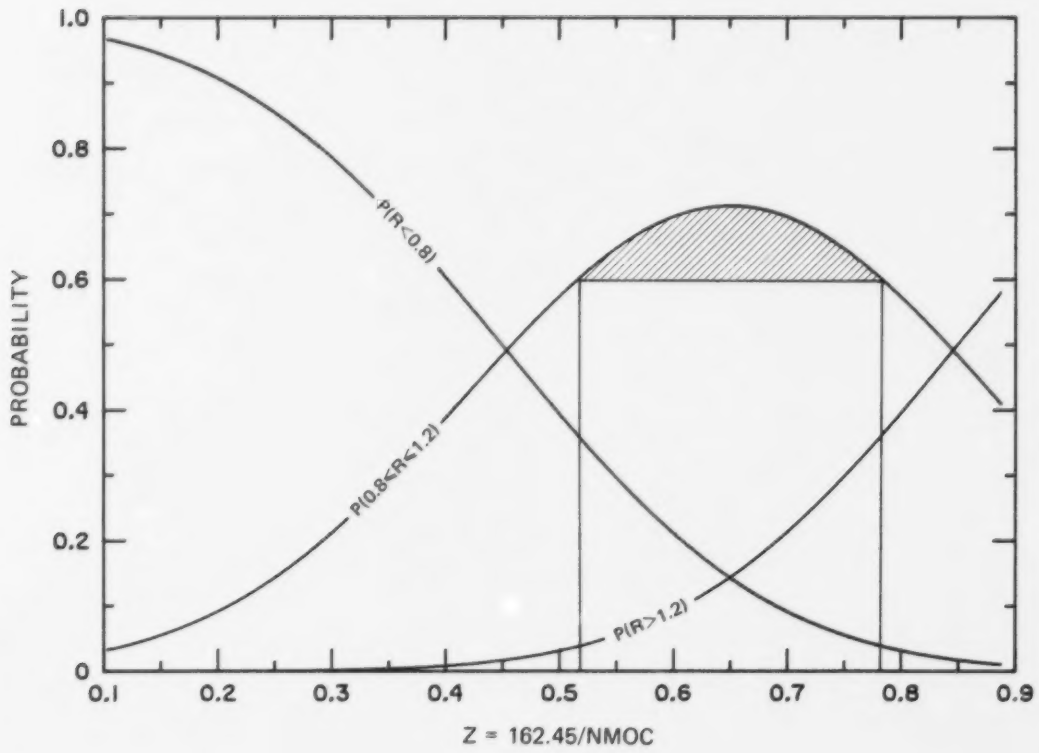


Figure 27. Accuracy probability plot for standard-EKMA ozone estimates for Philadelphia. Mean value is assumed for maximum daily temperature.

shaded area, the maximum probability is approximately 0.71, and occurs for Z = 0.65, which corresponds to NMOC = 250 ppbC. For the shaded area the variable Z is bounded by $0.51 < Z < 0.79$, for $P(0.8 \leq R \leq 1.2) > 0.6$.

The accuracy regions are illustrated in the NMOC-NO_x plane in Figure 28. The shaded area of the figure corresponds to that in Figure 27 and is defined by $206 < NMOC < 319$ ppbC. Thus, (NMOC, NO_x) combinations within the shaded area have the highest probability of yielding an accurate ozone estimate. Values of NMOC and NO_x to the right of the shaded area have a high probability that the ratio OBS/EST < 0.8; hence, this is the region of overprediction. Underprediction is most probable in the thin slice to the left of the shaded area. Thus, the vast majority of the evaluation region is associated with OBS/EST ratios smaller than 1.2.

EVALUATION OF CITY-SPECIFIC EKMA

City-Specific Ozone Estimates

Table 15 lists the input parameters used to calculate the city-specific ozone estimates for Philadelphia. As Table 13 showed, the highest observed ozone occurred on 13 July, which is the date used as the input for the model. The other parameters in Table 15 were obtained following the same procedures used in the cases previously described.

Continuing the pattern of the St. Louis and Houston data, the city-specific and standard-EKMA ozone estimates are strongly correlated, with r = 0.99 (see Table 14). The regression equation relating C-S and standard-EKMA ozone is CS03 = 36.01 + 0.804 * EST, with standard error s = 6.29, and units in ppb; the regression is highly statistically significant ($p < 0.00001$). The equation indicates that CS03 > EST for EST < 180 ppb. This range is greater than that for St. Louis, for which CS03 > EST for EST < 98 ppb.

Comparison of Observed Ozone and City-Specific Estimates

Figure 29 compares OBS and CS03 concentrations. The figure shows that Region 1 contains no points, Region 2 has eight points (28 percent) and Region 3 has 21 (72 percent). Thus, in contrast to the standard-EKMA estimates (cf. Figure 25), the C-S estimates include no underpredictions. Continuing the pattern of the standard-EKMA estimates, there is no statistically significant correlation between OBS and CS03.

Definition of Accuracy Regions

The multiple regression equation derived for the ratio OBS/CS03 is

$$OBS/CS03 = -0.580 + 117.72/NMOC + 0.0364(T) \quad . \qquad (9)$$

The multiple regression coefficient for Eq. 9 is r = 0.73, and the standard error is s = 0.167. The coefficients of 1/NMOC and T are statistically significant ($p < 0.025$), but the constant is not significant at the 0.05 level.

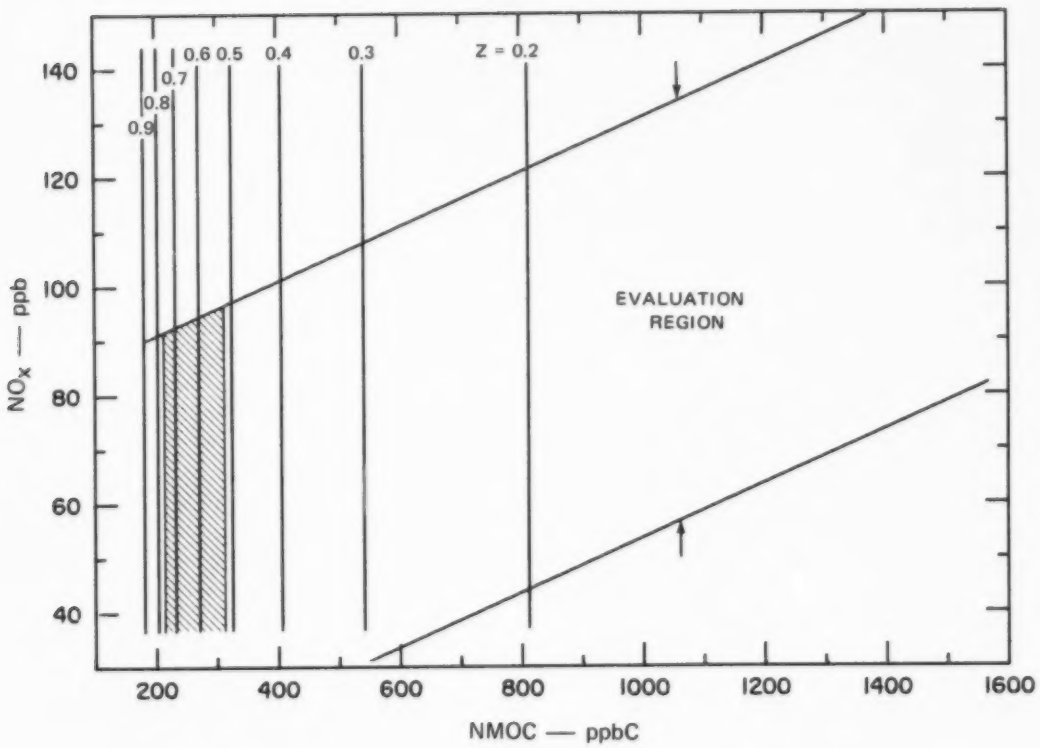


Figure 28. Constant-Z plot for standard-EKMA ozone estimates for Philadelphia. Shading defines area where ozone estimates have the highest probability of being accurate, assuming a maximum daily temperature of 27.5°C, which is its mean value.

TABLE 15. SUMMARY OF INPUT PARAMETERS FOR OBTAINING CITY-SPECIFIC EKMA OZONE ESTIMATES FOR PHILADELPHIA

Parameter	Value
Date	13 July 1979
Location	Latitude 40.00°N, Longitude 75.44°W
Inversion height data	
Initial height (m)	600
Final height (m)	1800
Start time of rise	0700
Ending time of rise	1400
Post-0800 emissions by ending hour	
0900	
NMOC	0.25
NO_x	0.25
1000	
NMOC	0.04
NO_x	0.04
1100	
NMOC	0.13
NO_x	0.13
1200	
NMOC	0.28
NO_x	0.28
1300	
NMOC	0.35
NO_x	0.35
Reactivity, NO_2/NO_x	0.58
Background ozone concentration (ppm)	0.070

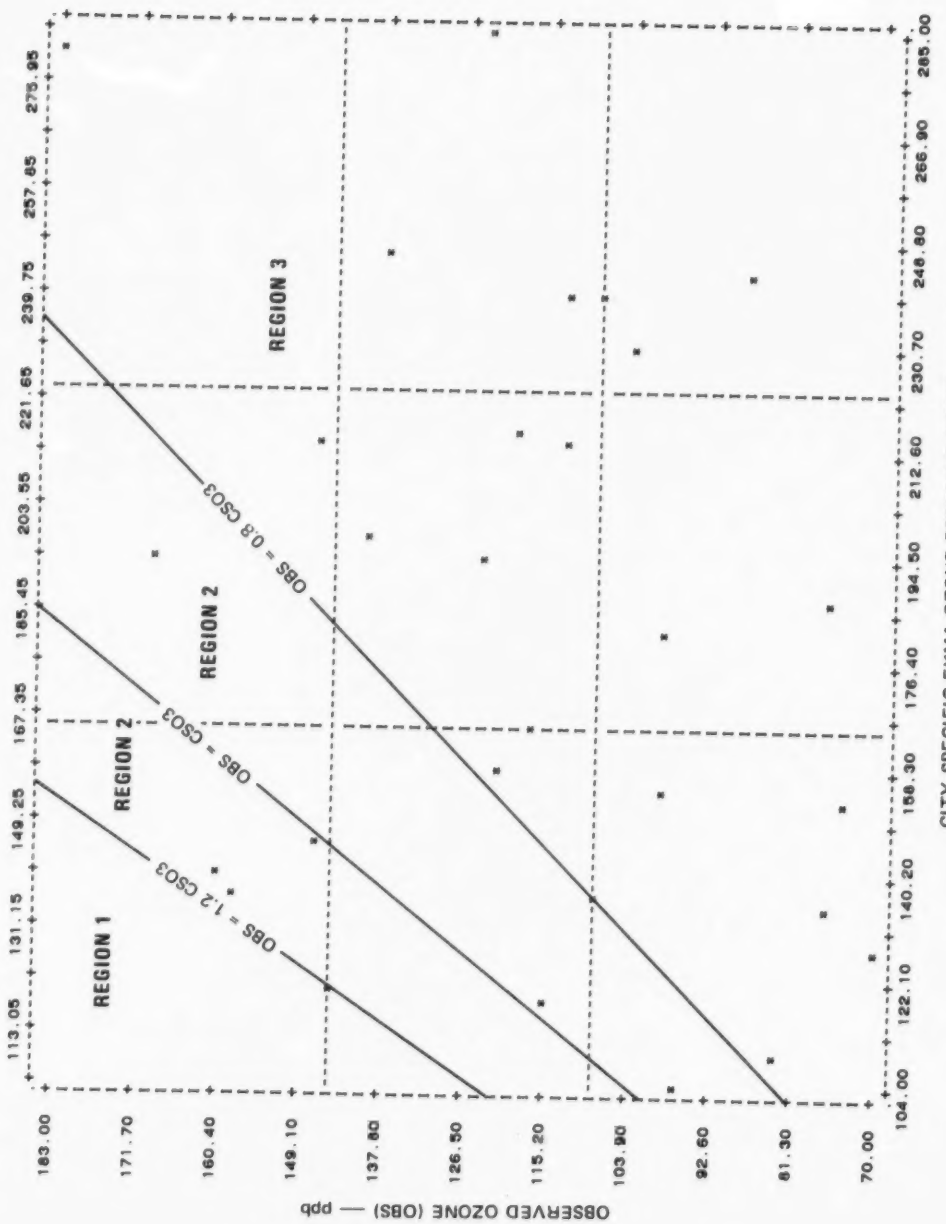


Figure 29. Scatterplot of observed ozone and city-specific EKMA ozone estimate for Philadelphia.

Equations 8 and 9 are very similar, which is a consequence of the strong correlation between EST and CS03. Moreover, the 95-percent confidence intervals of the coefficients of 1/NMOC and T for Eq. 8 overlap the corresponding intervals for Eq. 9.

Figure 30 shows the plot of the accuracy probabilities associated with Eq. 9 for T = 27.5°C, which is its mean value. The graph shows probability as a function of the variable Z = 117.72/NMOC. The region where $P(R \leq 1.2) > 0.5$ is defined by $Z < 0.78$. In this region, the ozone estimate has a better than 50-percent probability of yielding an upper bound to the observed ozone. The shaded area in Figure 30 marks the region where $P(0.8 \leq R \leq 1.2) > 0.6$. Values of Z within the shaded area are defined by $0.43 < Z < 0.73$, and correspond to 161 < NMOC < 274 ppbC. Thus, the C-S ozone estimates are most likely to be accurate where NMOC is within these bounds.

The accuracy regions are displayed on the NMOC-NO_x plane in Figure 31. The shaded area of this figure corresponds to that of Figure 30; here C-S ozone estimates have the highest probability of being accurate. The upper-bound region, i.e. were $R \leq 1.2$, includes the shaded area and all of the evaluation region to the right of the shaded area. Only a very narrow slice of the NMOC-NO_x plane to the left of the shaded area corresponds to the region of underprediction, and that slice is actually outside the bounds of the evaluation region. This is, of course, a reflection of the fact that no points fell in Region 1 of Figure 29.

DISCUSSION

For Philadelphia, the standard- and city-specific EKMA ozone estimates were very similar. However, in a reversal of roles from the St. Louis and Houston cases, the Philadelphia C-S estimates were more accurate and displayed a lower tendency toward underprediction than did the standard-EKMA estimates. However, in keeping with previous results, the standard-EKMA ozone estimates showed a pronounced tendency toward overprediction. Because the city-specific EKMA yielded more accurate estimates, and because its upper-bound properties are similar to the standard EKMA's, it is recommended that the city-specific EKMA, rather than the standard EKMA, should be the method used for application to the Philadelphia area. This contrasts with the Houston situation, in which the underpredictive tendency of the city-specific EKMA renders it useless for an upper-bound type of analysis.

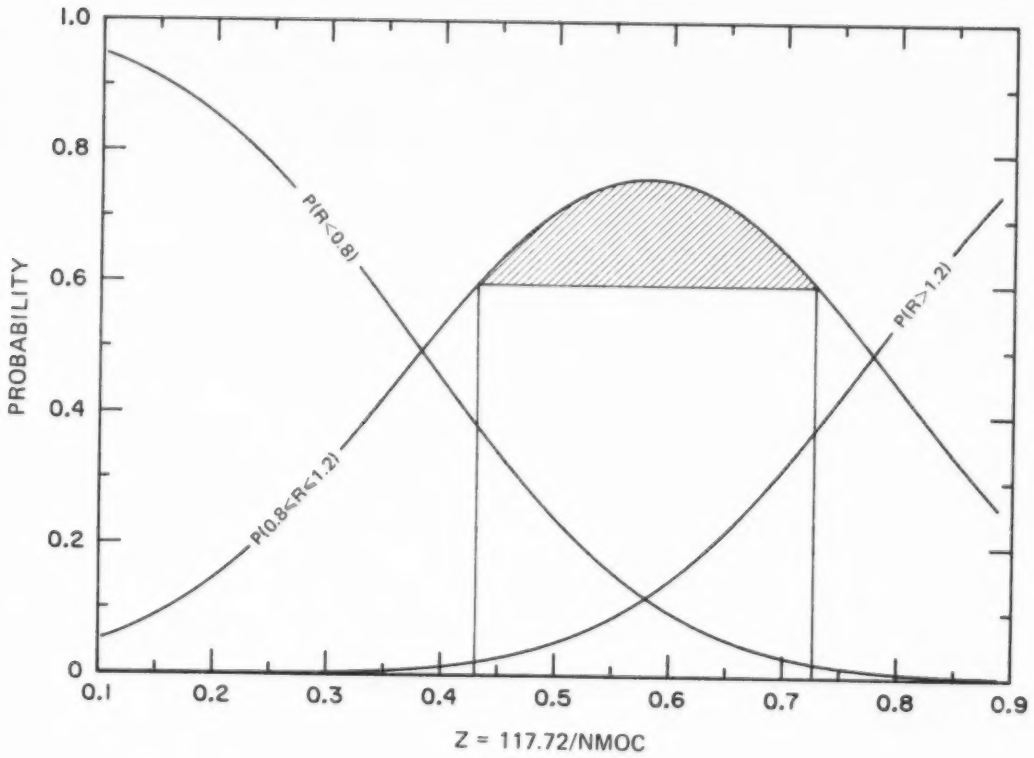


Figure 30. Accuracy probability plot for city-specific ozone estimates for Philadelphia. Mean value is assumed for maximum daily temperature.

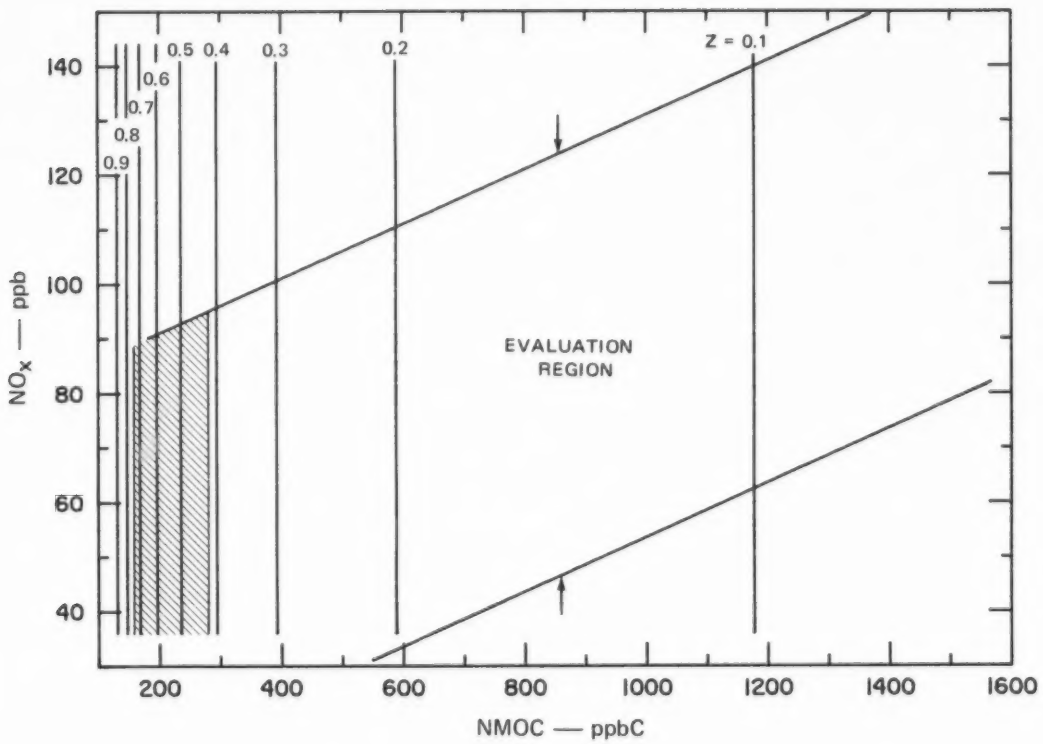


Figure 31. Constant-Z plot for city-specific EKMA ozone estimates for Philadelphia. Shaded area denotes region where ozone estimates have the highest probability of being accurate, assuming the mean value of 27.5°C for the maximum daily temperature.

SECTION 6

EVALUATION USING DATA FOR THE LOS ANGELES AREA

DATA REVIEW AND ANALYSIS

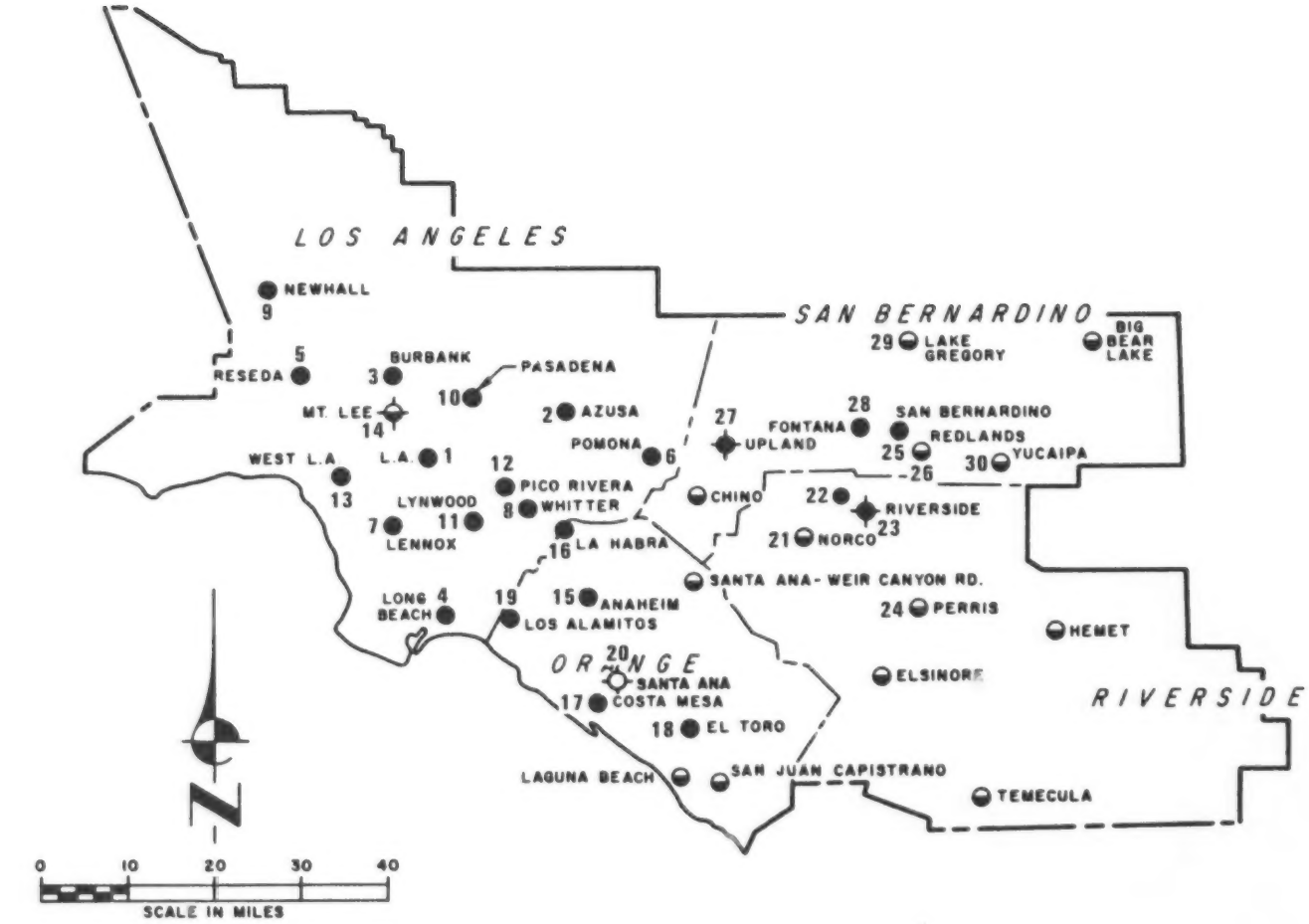
The EKMA was evaluated using data from the South Coast Air Basin (the Los Angeles Basin) air monitoring network, operated by the California Air Resources Board and the South Coast Air Quality Management District. The data were collected at thirty monitoring stations in the South Coast Air Basin (Figure 32) during the period May through October 1978. Data for 1976 and 1977 were also examined; 1978 was chosen because it had the highest ozone levels.

From a total of 184 days, only 20 days did not have observed ozone concentrations exceeding the NAAQS for ozone of 120 ppb. The highest observed hourly averaged ozone concentration was 460 ppb. The daily maximum hourly ozone concentrations were significantly correlated with daily maximum temperature; however, the correlation coefficient was low ($r = 0.24$). As with the Houston data, the low correlation between peak ozone and peak temperature is probably due to the narrow range in daily maximum temperatures. The daily maximum ozone was also correlated with the 0600-0900 average source region NMOC and NO_x concentrations, but the correlation was low, although statistically significant.

Six monitoring sites were used to calculate the 0600-0900 average source region NMOC and NO_x concentrations:

- Site 1--downtown Los Angeles
- Site 3--Burbank
- Site 4--Long Beach
- Site 6--Pomona
- Site 7--Lennox
- Site 8--Whittier.

Results from an analysis of variance performed on the 0600-0900 average NMOC and NO_x concentrations monitored at each of these six sites showed that the number of sites could not be condensed on a statistical basis.



SOURCE: California Air Resources Board.

Figure 32. Air-quality monitoring network for the Los Angeles area.

DEFINITION OF EVALUATION DATA SET

No selection criteria were used to decrease the number of test days selected from the Los Angeles data base. Because the peak ozone concentration exceeded the 120 ppb NAAQS for ozone on most days, the only days excluded from the data base were those that had the 0600-0900 average source region NMOC or NO_x concentrations missing. Two additional days (14 September and 30 October) were excluded because the daily maximum ozone concentrations (90 ppb and 80 ppb) occurred at 0300 and 1000, respectively.

The evaluation data set is listed in Table 16. Included in the table are the daily 0600-0900 average source region core city NMOC and NO_x concentrations; daily maximum temperature; the site, time, and magnitude of the observed daily maximum ozone concentration; and the standard-EKMA and city-specific ozone estimates. The number of test days is 176.

EVALUATION OF STANDARD EKMA

Comparison of Observed and Estimated Ozone

Figure 33 is a scatterplot of OBS and EST for Los Angeles. As the scatter suggests, no statistically significant correlation exists between observation and prediction. Nevertheless, the distribution of the points among the three accuracy regions is of interest. Region 1 contains 62 points (35 percent), Region has 51 (29 percent), and Region 3 has 63 (36 percent). Thus, overpredictions, underpredictions, and accurate predictions are about equally probable. This is surprising, because the standard-EKMA is supposed to simulate worst-case ozone conditions in the Los Angeles area. However, the figure shows that 35 percent of the cases are underpredicted, which implies that the standard-EKMA is not useful for obtaining upper bounds for maximum ozone potential in the Los Angeles area.

Relationships Between the Variables

Table 17 shows correlation coefficients for selected variables in the Los Angeles data set. Both standard-EKMA and city-specific ozone estimates are included in the table, but the latter will be discussed in another section. The table shows that OBS is weakly correlated with 1/NMOC and $1/NO_x$. By contrast, EST is strongly correlated with NMOC, NO_x, 1/NMOC and $1/NO_x$. and, as a consequence, the ratio OBS/EST is also correlated with these variables. Maximum temperature is also correlated with NMOC, NO_x, 1/NMOC, $1/NO_x$. and EST, which contrasts with the situation in St. Louis, where no significant correlation existed.

Table 17 indicates that NMOC and NO_x are strongly correlated. This is evident in Figure 34, which displays the strong linear relationship between the two variables. The correlation coefficient is $r = 0.92$, and the regression line (which is plotted on the graph) is $NO_x = 41.68 + 0.142$ (NMOC), where the units are as defined in Table 17. The standard error of the regression is $s = 29.96$ ppb, which was used in conjunction with the regression line to define the evaluation region in Figure 34.

TABLE 16. EVALUATION DATA SET FOR LOS ANGELES

DATE (1978)	NMOC (ppbC)	NOX (ppb)	TEMP (DEG F)	OBSERVED OZONE			EKMA OZONE ESTIMATES (ppb)	
				PEAK STN	PEAK HOUR	PPB	STANDARD	CITY-SPECIFIC
1 MAY	373	123	66	19	15	80	125	227
2 MAY	779	171	68	28	15	130	262	516
3 MAY	663	130	62	26	15	180	242	280
4 MAY	299	89	65	21	15	190	119	237
5 MAY	350	84	66	22	15	100	151	312
6 MAY	1050	156	70	28	18	130	292	589
7 MAY	828	130	76	29	18	120	251	530
8 MAY	1392	296	81	29	17	210	389	695
9 MAY	1470	231	77	28	14	190	370	699
10 MAY	425	109	68	28	14	110	165	338
11 MAY	634	103	68	28	16	190	211	460
12 MAY	1773	265	85	21	16	400	417	745
13 MAY	1589	179	82	29	17	280	355	631
14 MAY	868	123	80	29	14	190	254	527
15 MAY	411	78	72	28	15	110	177	362
16 MAY	406	97	75	16	15	120	166	343
17 MAY	1369	284	84	22	18	110	387	694
18 MAY	1392	273	86	29	18	170	395	704
19 MAY	1531	260	74	14	14	260	384	726
20 MAY	523	95	68	29	15	220	187	418
21 MAY	303	52	68	25	12	170	129	295
22 MAY	496	71	69	29	15	140	175	382
23 MAY	373	70	68	22	14	80	167	340
24 MAY	871	179	67	22	15	100	287	556
25 MAY	902	231	69	30	16	120	271	500
26 MAY	649	172	70	22	15	130	214	414

TABLE 16 (continued)

DATE (1978)	NMOC (ppbC)	NOX (ppb)	TEMP (DEG F)	OBSERVED OZONE			EKMA OZONE ESTIMATES (ppb)	
				PEAK STN	PEAK HOUR	PPB	STANDARD	CITY-SPECIFIC
27 MAY	1017	185	73	25	16	240	292	603
28 MAY	1297	167	78	15	13	230	327	614
29 MAY	1517	164	95	16	13	260	346	604
31 MAY	187	72	71	28	15	150	67	120
1 JUN	257	68	69	30	17	120	118	243
2 JUN	484	63	71	26	17	160	169	367
3 JUN	638	113	70	28	16	230	213	466
4 JUN	233	87	70	9	15	250	80	141
5 JUN	420	98	71	9	16	300	171	353
6 JUN	457	101	71	2	14	370	185	380
7 JUN	420	107	72	2	14	370	165	336
8 JUN	523	91	73	2	13	420	186	417
9 JUN	868	134	78	2	15	390	258	541
10 JUN	350	50	72	9	15	250	139	310
12 JUN	1027	190	79	2	16	320	324	607
13 JUN	1408	242	80	10	14	300	363	701
14 JUN	684	128	75	10	13	190	248	489
15 JUN	523	122	76	3	13	180	198	402
16 JUN	1260	195	74	28	15	320	332	648
17 JUN	1143	201	79	2	14	390	316	638
18 JUN	968	144	76	10	12	230	277	560
19 JUN	1246	271	77	26	16	320	359	659
20 JUN	1176	227	78	28	16	360	354	650
21 JUN	915	167	77	2	15	350	300	572
22 JUN	943	158	83	9	14	320	276	576
23 JUN	1755	230	90	2	15	460	492	704

TABLE 16 (continued)

DATE (1978)	NMOC (ppbC)	NOX (ppb)	TEMP (DEG F)	OBSERVED OZONE			EKMA OZONE ESTIMATES (ppb)	
				PEAK STN	PEAK HOUR	PPB	STANDARD	CITY-SPECIFIC
24 JUN	1335	193	76	2	12	310	342	654
25 JUN	572	95	76	9	15	200	197	433
26 JUN	303	76	75	9	13	160	135	278
29 JUN	327	106	71	10	13	170	117	220
30 JUN	613	160	74	9	14	220	207	409
1 JUL	591	110	73	10	14	200	224	449
2 JUL	513	91	73	2	15	200	184	412
3 JUL	315	101	73	20	15	200	115	220
4 JUL	206	80	74	9	15	200	113	221
5 JUL	373	130	73	9	15	240	117	203
6 JUL	624	165	76	20	15	290	200	409
7 JUL	525	95	72	20	15	320	187	417
8 JUL	933	150	73	20	15	370	274	575
9 JUL	455	86	69	20	15	300	190	384
10 JUL	389	103	72	2	14	210	153	313
11 JUL	373	90	71	2	16	250	157	323
12 JUL	669	163	74	2	14	310	220	451
13 JUL	1073	200	83	6	16	450	334	621
14 JUL	1100	249	83	2	14	410	329	620
15 JUL	992	160	85	2	14	450	285	587
16 JUL	397	82	82	2	14	290	171	351
17 JUL	824	138	75	2	13	230	252	536
18 JUL	770	157	72	25	16	270	265	521
19 JUL	560	117	70	2	14	300	214	433
20 JUL	747	173	70	9	15	300	250	495
21 JUL	863	174	72	20	14	300	286	554

TABLE 16 (Continued)

DATE (1978)	NMOC (ppbC)	NOX (ppb)	TEMP (DEG F)	OBSERVED OZONE			EKMA OZONE ESTIMATES (ppb)	
				PEAK STN	PEAK HOUR	PPB	STANDARD	CITY-SPECIFIC
22 JUL	852	171	73	2	13	260	284	549
23 JUL	741	108	70	28	15	360	230	489
24 JUL	688	161	73	28	16	330	237	469
25 JUL	601	138	74	2	14	300	210	441
26 JUL	945	173	77	2	14	330	307	580
27 JUL	747	148	70	9	14	310	261	513
28 JUL	848	107	76	28	13	300	277	540
29 JUL	420	140	74	2	13	320	132	235
30 JUL	342	75	75	2	13	290	153	310
31 JUL	630	130	75	9	13	280	232	462
1 AUG	411	115	77	28	15	280	153	305
2 AUG	495	115	75	9	14	240	191	392
3 AUG	476	139	75	29	16	320	163	315
4 AUG	471	126	76	12	13	360	173	347
5 AUG	448	139	81	6	14	390	149	277
6 AUG	448	106	80	3	15	270	170	365
7 AUG	502	130	76	2	13	320	176	340
8 AUG	336	103	79	29	15	290	125	244
9 AUG	476	125	81	28	15	230	176	354
10 AUG	467	110	81	29	16	290	103	373
11 AUG	303	97	75	29	10	230	113	216
12 AUG	230	69	75	2	14	220	105	214
14 AUG	303	95	77	2	13	100	115	223
15 AUG	485	127	74	2	14	230	170	350
16 AUG	933	176	72	26	12	190	304	579
17 AUG	602	150	72	2	13	170	210	420

TABLE 16 (continued)

DATE (1978)	NMOC (ppbC)	NOX (ppb)	TEMP (DEG F)	OBSERVED OZONE			EKMA OZONE ESTIMATES (ppb)	
				PEAK STN	PEAK HOUR	PPB	STANDARD	CITY-SPECIFIC
18 AUG	1000	222	74	29	17	250	311	591
19 AUG	560	121	73	2	15	270	212	429
20 AUG	513	100	74	2	15	230	205	414
21 AUG	723	155	74	29	15	180	251	498
22 AUG	572	155	72	2	14	220	194	380
23 AUG	672	142	77	10	14	160	240	482
24 AUG	1241	247	75	10	15	190	365	667
25 AUG	905	207	75	29	16	210	286	549
26 AUG	831	155	75	2	14	240	202	544
27 AUG	887	121	76	25	15	260	256	526
28 AUG	1269	250	82	28	15	340	371	674
29 AUG	443	142	77	9	15	410	143	260
30 AUG	436	93	79	9	14	310	101	372
31 AUG	728	128	76	2	15	350	233	505
1 SEP	1027	100	77	10	14	340	294	604
2 SEP	583	111	74	2	14	370	223	446
3 SEP	863	124	79	2	14	280	254	522
4 SEP	752	126	80	10	15	190	237	500
5 SEP	625	140	77	29	13	70	226	453
6 SEP	330	77	73	9	14	70	151	313
7 SEP	513	139	75	10	15	170	181	359
8 SEP	1409	276	77	9	15	260	390	707
9 SEP	653	103	75	2	15	170	214	465
10 SEP	875	113	75	2	14	170	251	506
11 SEP	714	154	77	27	15	170	249	492
12 SEP	840	161	77	9	15	220	283	546

TABLE 16 (continued)

DATE (1978)	NMOC (ppbC)	NOX (ppb)	TEMP (DEG F)	OBSERVED OZONE			EKMA OZONE ESTIMATES (ppb)	
				PEAK STN	PEAK HOUR	PPB	STANDARD	CITY-SPECIFIC
13 SEP	552	108	71	30	15	120	214	431
15 SEP	1091	240	73	27	15	150	328	616
16 SEP	369	106	73	10	13	200	140	279
18 SEP	1008	249	75	26	17	100	297	553
19 SEP	1960	427	88	12	15	90	486	798
20 SEP	1997	397	91	12	17	110	502	821
21 SEP	1855	349	93	2	15	130	400	798
22 SEP	2049	371	92	7	16	260	469	829
23 SEP	1372	236	100	17	13	220	357	694
24 SEP	1202	167	105	7	15	290	316	616
25 SEP	2520	352	105	12	14	230	522	829
26 SEP	2427	388	97	10	14	360	521	859
27 SEP	1980	304	87	9	15	320	452	786
28 SEP	1057	254	87	12	14	300	423	734
29 SEP	1661	310	84	2	14	430	446	762
30 SEP	961	177	84	28	16	200	310	587
1 OCT	443	67	76	9	15	270	164	367
2 OCT	200	71	76	9	14	200	127	263
3 OCT	566	152	74	9	15	300	194	382
4 OCT	507	146	73	25	16	330	172	330
5 OCT	577	117	74	2	15	300	220	441
6 OCT	776	165	74	9	16	390	264	518
7 OCT	817	121	76	9	15	390	247	518
8 OCT	315	121	78	2	14	200	92	149
9 OCT	1405	264	77	2	15	100	399	700
10 OCT	1043	342	78	21	15	210	479	797

TABLE 16 (concluded)

DATE (1978)	NMOC (ppbC)	NOX (ppb)	TEMP (DEG F)	OBSERVED OZONE			EKMA OZONE ESTIMATES (ppb)	
				PEAK STN	PEAK HOUR	PPB	STANDARD	CITY-SPECIFIC
11 OCT	1848	308	88	2	14	270	435	782
12 OCT	2123	385	78	10	15	300	467	788
13 OCT	1657	253	74	6	16	400	399	738
14 OCT	1204	212	74	12	14	390	327	655
15 OCT	1232	197	72	2	15	280	329	649
16 OCT	989	182	75	9	14	220	316	595
17 OCT	303	98	75	9	14	120	112	212
18 OCT	1148	217	72	8	13	170	349	643
19 OCT	1988	326	76	21	15	150	456	803
20 OCT	723	135	74	9	14	110	257	505
21 OCT	835	135	72	9	15	160	253	535
22 OCT	758	80	73	2	14	150	215	433
23 OCT	714	145	89	12	13	70	252	498
24 OCT	2345	335	77	26	15	80	499	816
25 OCT	665	128	72	9	15	120	243	481
26 OCT	1475	308	71	9	15	180	406	715
27 OCT	1213	238	69	9	15	190	360	659
28 OCT	898	183	69	2	15	240	293	564
29 OCT	548	80	69	10	15	240	108	412
31 OCT	1050	206	67	9	14	80	329	617

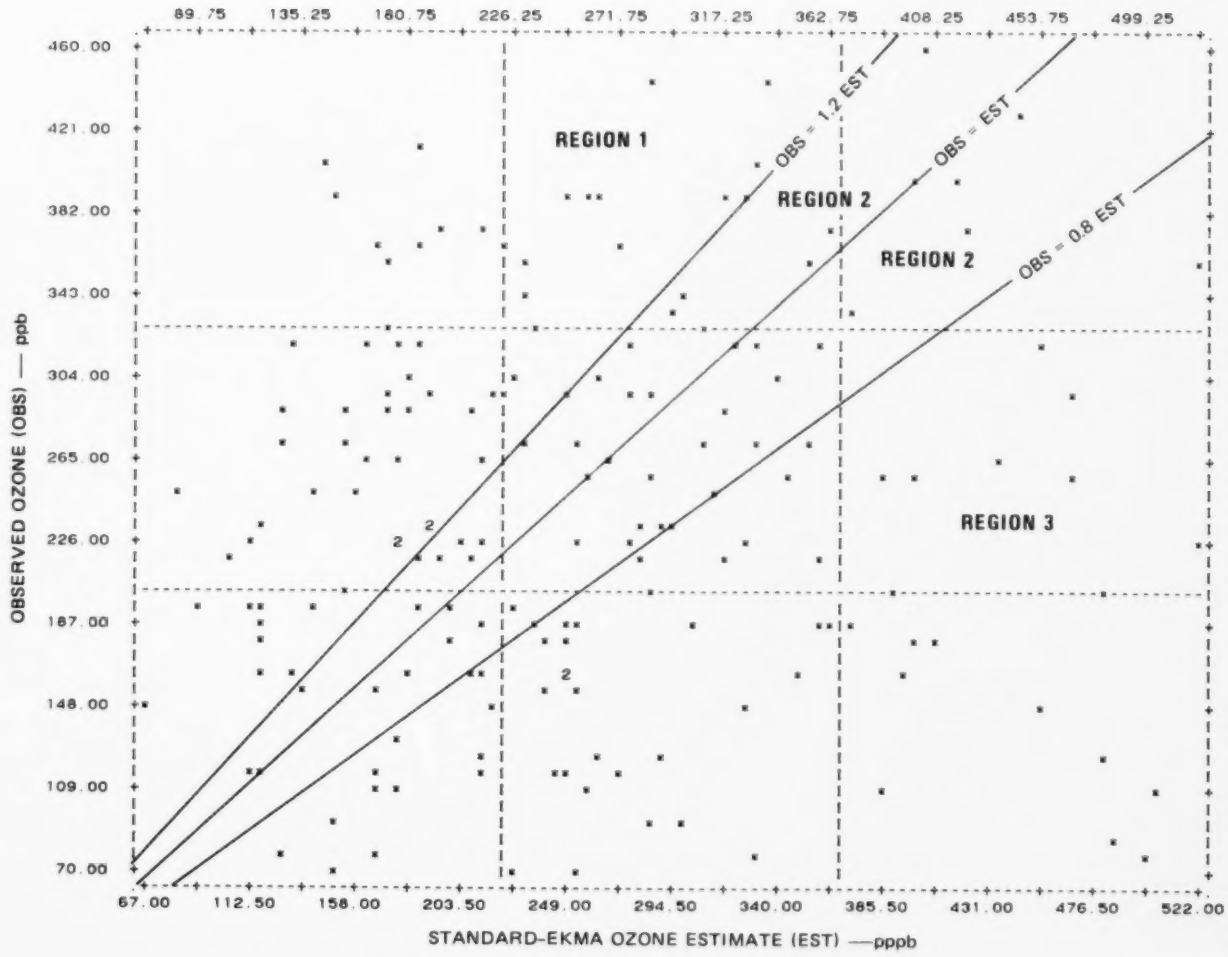


Figure 33. Scatterplot of observed ozone and standard-EKMA ozone estimate for Los Angeles.

TABLE 17. MEANS, STANDARD DEVIATIONS, AND CORRELATIONS FOR SELECTED VARIABLES IN THE LOS ANGELES DATA SET

Variable	Mean	Standard Deviation	Correlation Coefficient*								
			NO_x	1/NMOC	$1/NO_x$	Maximum Temperature	OBS	EST	OBS/EST	CSO3	OBS/CSO3
NMOC (ppbC)	854.4	496.5	0.92	-0.82	-0.78	0.59	--	0.98	-0.57	0.94	-0.51
NO_x (ppb)	162.8	76.6		-0.75	-0.86	0.52	--	0.92	-0.53	0.87	-0.45
1/NMOC ($ppbC^{-1}$)	0.0016	0.0009			0.82	-0.42	-0.19†	-0.89	0.63	-0.93	0.61
$1/NO_x$ (ppb^{-1})	0.0075	0.003				-0.43	-0.18†	-0.81	0.46	-0.79	0.39
Temperature (^oF)	76.1	6.8					0.23	0.54	-0.16§	0.50	--
OBS (ppb)‡	243.9	93.7						--	0.56	0.14**	0.56
EST (ppb)††	256.6	102.1							-0.64	0.98	0.99
OBS/EST	1.10	0.57								-0.66	-0.63
CSO3 (ppb)‡‡	493.3	168.7									
OBS/CSO3	0.56	0.30									1.0

*All correlations significant at the 0.001 level unless noted otherwise. Dashes indicate that correlation is not significant at the 0.05 level.

†Significant level < 0.01.

‡Observed ozone.

§Significant level = 0.018.

**Significant level < 0.03.

††Standard-EKMA ozone estimate.

‡‡City-specific EKMA ozone estimate.

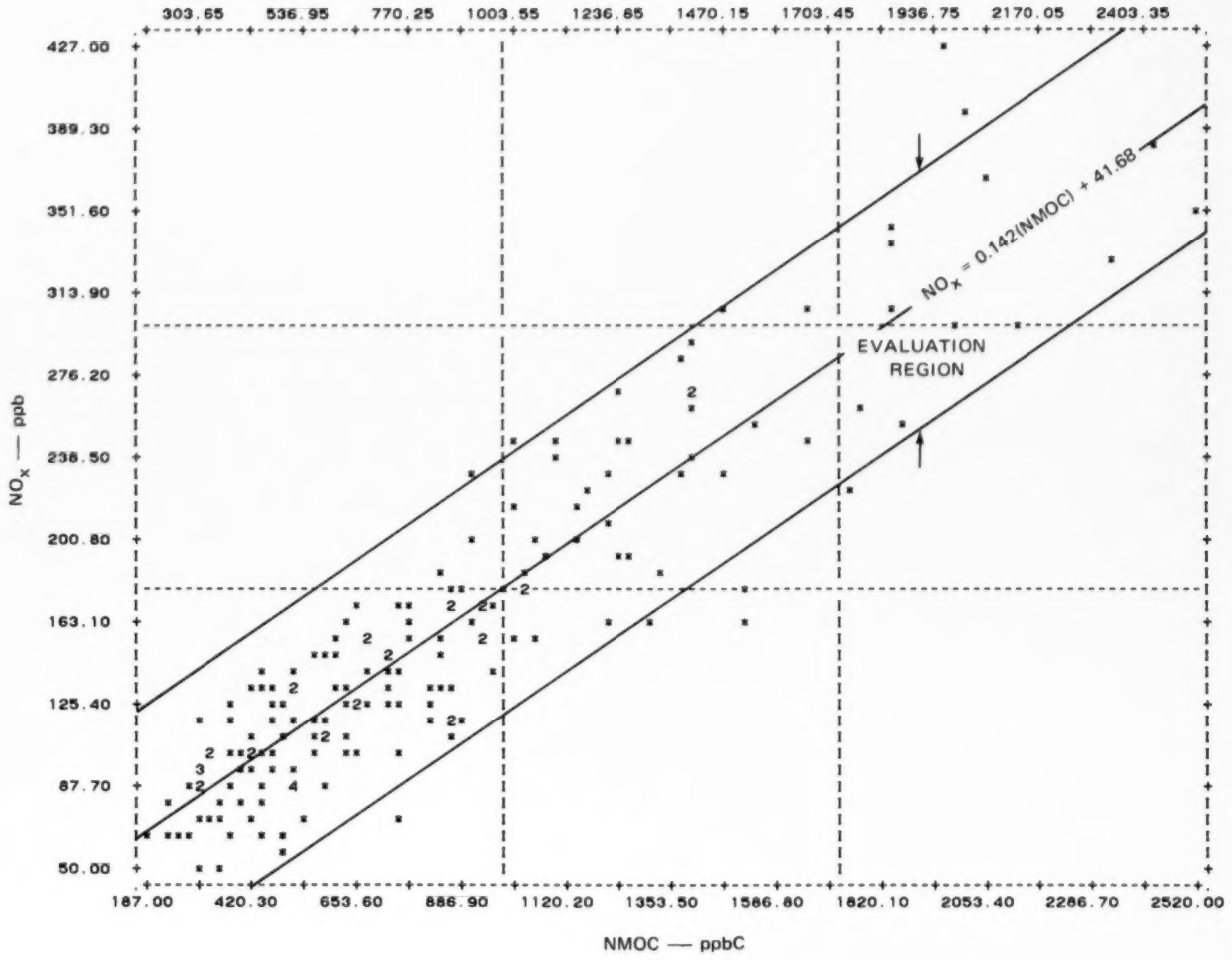


Figure 34. Scatterplot of 0600–0900 NMOC and NO_x for Los Angeles.

Definition of Accuracy Regions

A multiple regression equation was derived for the ratio OBS/EST as a function of several variables:

$$OBS/EST = -0.244 - 39.06/NO_x + 371.60/NMOC \quad (10)$$

$$-0.000456(NMOC) + 0.0187(T) \quad ,$$

where T denotes daily maximum temperature in ^{o}F and the other symbols and units are as previously defined. The multiple regression coefficient is $r = 0.68$, and the standard error of the regression is $s = 0.43$. All the coefficients of Eq. 10 are statistically significant at better than the 0.05 level except the constant, which is not.

Equation 10 resembles Eq. 4 for St. Louis because it includes both 1/NMOC and $1/NO_x$. However, in Eq. 4 these two variables had positive coefficients, which is not the case in Eq. 10. Another difference between Eq. 4 and 10 is that the latter includes NMOC; despite the fact that 1/NMOC, $1/NO_x$, and NMOC are intercorrelated (see Table 17) the latter two variables increase the amount of variance explained. By far the most important variable in terms of amount of variances explained is 1/NMOC, which by itself explains about 40 percent of the variance. The remaining three variables--T, NMOC and $1/NO_x$--together add another 6 percent to the total explained variance. Thus, NMOC or its reciprocal continues to play a large role in explaining the predictive performance of the EKMA.

Figure 35 displays a plot of accuracy probability derived from Eq. 10 after setting $T = 76.1^{o}F$, which is its mean value. The variable $Z = -39.06/NO_x + 371.60/NMOC - 0.000456$ (NMOC) is the right side of Eq. 10 without the constant and the factor for T. Reflecting the indications of Figure 33, Figure 35 shows that the probability of an accurate prediction is relatively low, with $P(0.8 \leq R \leq 1.2) < 0.37$. Moreover, the magnitude of the three probabilities is about the same in the neighborhood where $P(0.8 \leq R \leq 1.2)$ has its maximum. Thus, relatively small changes in the value of Z in this neighborhood can radically shift the probability of an accurate prediction in favor of an increased probability of overprediction or underprediction. The steepness of the curves for $P(R < 0.8)$ and for $P(R > 1.2)$ suggests a similarly sensitive behavior in the overprediction and underprediction regimes, respectively. The sensitivity of the probabilities causes the standard-EKMA ozone estimates to be of limited usefulness for obtaining upper bounds for ozone in Los Angeles.

Figure 36 is the companion to Figure 35, depicting constant-Z contours in the NMOC-NO_x plane. For mean daily maximum temperature, the area to the right of the curve $Z = -0.50$ represents the region where overprediction is most probable. Conversely, the area to the left of the curve $Z = 0.25$ denotes the region of underprediction. The area between $Z = -0.5$ and $Z = 0.25$ may be considered a transition region where the three possibilities, i.e. overprediction, accurate prediction, and underprediction, have similar probabilities.

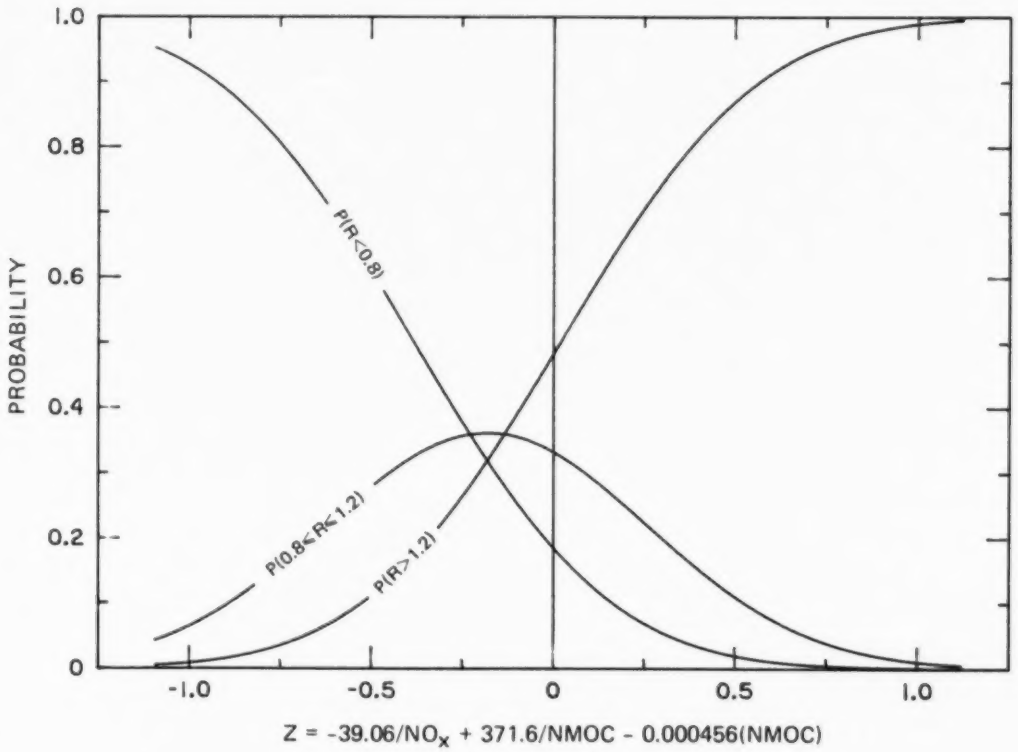


Figure 35. Accuracy probability plot for standard-EKMA ozone estimates for Los Angeles. Mean value is assumed for maximum daily temperature.

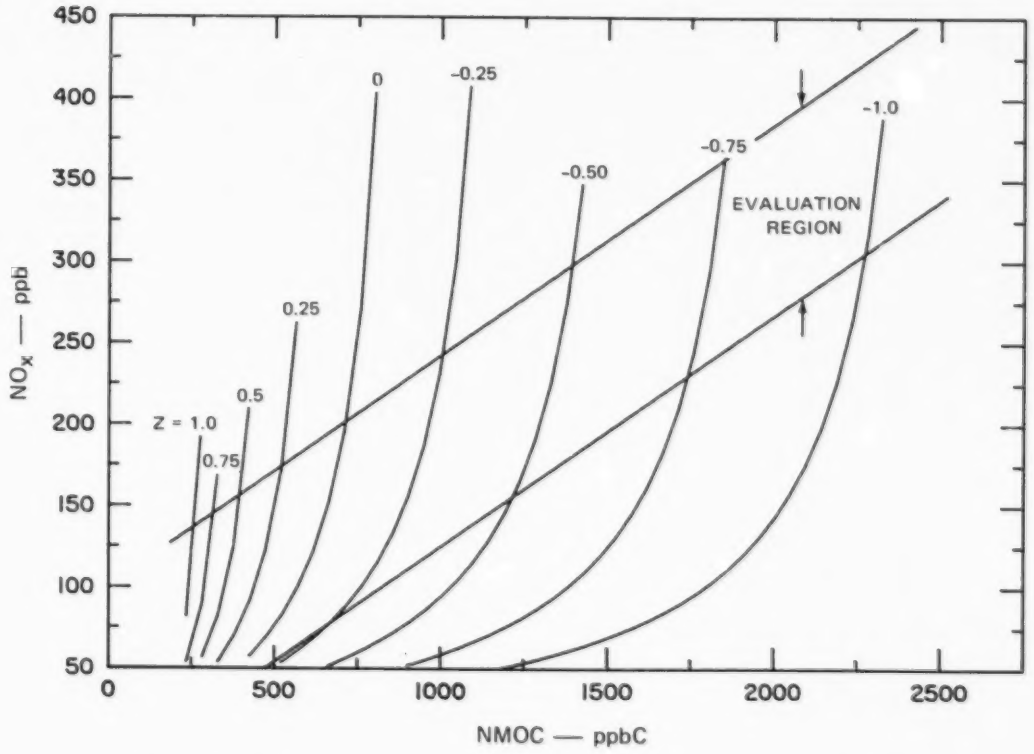


Figure 36. Plot of constant-Z curves on NMOC-NO_x plane for standard-EKMA ozone estimates for Los Angeles.

EVALUATION OF CITY-SPECIFIC EKMA

City-Specific Ozone Estimates

Table 18 shows the input parameters used with the EKMA computer program to obtain the C-S ozone estimates. As with the others, the date used, 23 June, was when the highest ozone level, 460 ppb, was observed. (This is one day after the summer solstice.) Default values were used for the inversion height; they are the same as for the standard EKMA. The city-specific EKMA differs from the standard EKMA mainly in that the former has post-0800 emissions, and the NO_2/NO_x ratio is 0.41 rather than the standard value of 0.25.

As in the other test cities the city-specific EKMA ozone estimates for Los Angeles, are highly correlated with the standard-EKMA estimates. Figure 37 shows a scatterplot of CS03 and EST that illustrates the strong relationship between the two variables. The correlation coefficient is $p = 0.98$, and is significant with $p < 0.00001$. The regression equation relating the two variables is CS03 = 78.55 + 1.62(EST), and has a standard error $s = 30.31$, where units are ppb. Actually, the figure indicates a nonlinear relationship exists between CS03 and EST, as evidenced by the slight curvature of the point swarm.

Like Philadelphia, but unlike St. Louis and Houston, the Los Angeles C-S estimates are greater than the Los Angeles standard-EKMA estimates. This is reflected in the intercept and slope of the regression equation, the slope indicating a scaling factor of 162 percent. This will lead to fewer underpredictions and to more overpredictions.

Comparison of Observed Ozone and City-Specific Ozone Estimate

Figure 38 shows a scatterplot of OBS and CS03. In contrast to Figure 33, Figure 38 has a preponderance of overpredictions and few underpredictions. Region 1 contains six points (3 percent of the total), Region 2 has 25 (14 percent), and Region 3 has 145 (83 percent). Thus, the number of underpredictions has been reduced by more than a factor of ten, but the number of accurate predictions has decreased by a factor of two. Although the magnitude of the overpredictions can be very large, more than a factor of ten in some cases, for purposes of estimating upper bounds the C-S EKMA is more appropriate then the standard-EKMA because of its low probability of underprediction.

Definition of Accuracy Regions

The multiple regression equation derived for the ratio OBS/CS03 is

$$OBS/CS03 = 0.0756 - 54.999/NO_x + 286.822/NMOC$$

$$-0.001687(NO_x) + 0.00932(T) \qquad (11)$$

where symbols and units are as defined for Eq. 10. The multiple regression coefficient is $r = 0.69$, and the standard error of the regression is $s = 0.22$.

TABLE 18. SUMMARY OF INPUT PARAMETERS FOR OBTAINING CITY-SPECIFIC EKMA OZONE ESTIMATES FOR LOS ANGELES

Parameter	Value
Date	23 June 1978
Location	Latitude 24.00°N, Longitude 118.5°W
Inversion data	Default values used
Post-0800 emissions by ending hour	
0900	
NMHC	0.61
NO_x	0.64
1000	
NMHC	0.57
NO_x	0.51
1100	
NMHC	0.35
NO_x	0.34
1200	
NMHC	0.30
NO_x	0.29
1300	
NMHC	0.25
NO_x	0.25
1400	
NMHC	0.16
NO_x	0.15
1500	
NMHC	0.14
NO_x	0.14
Reactivity	
NO_2/NO_x	0.41
Background ozone concentration (ppm)	0.04

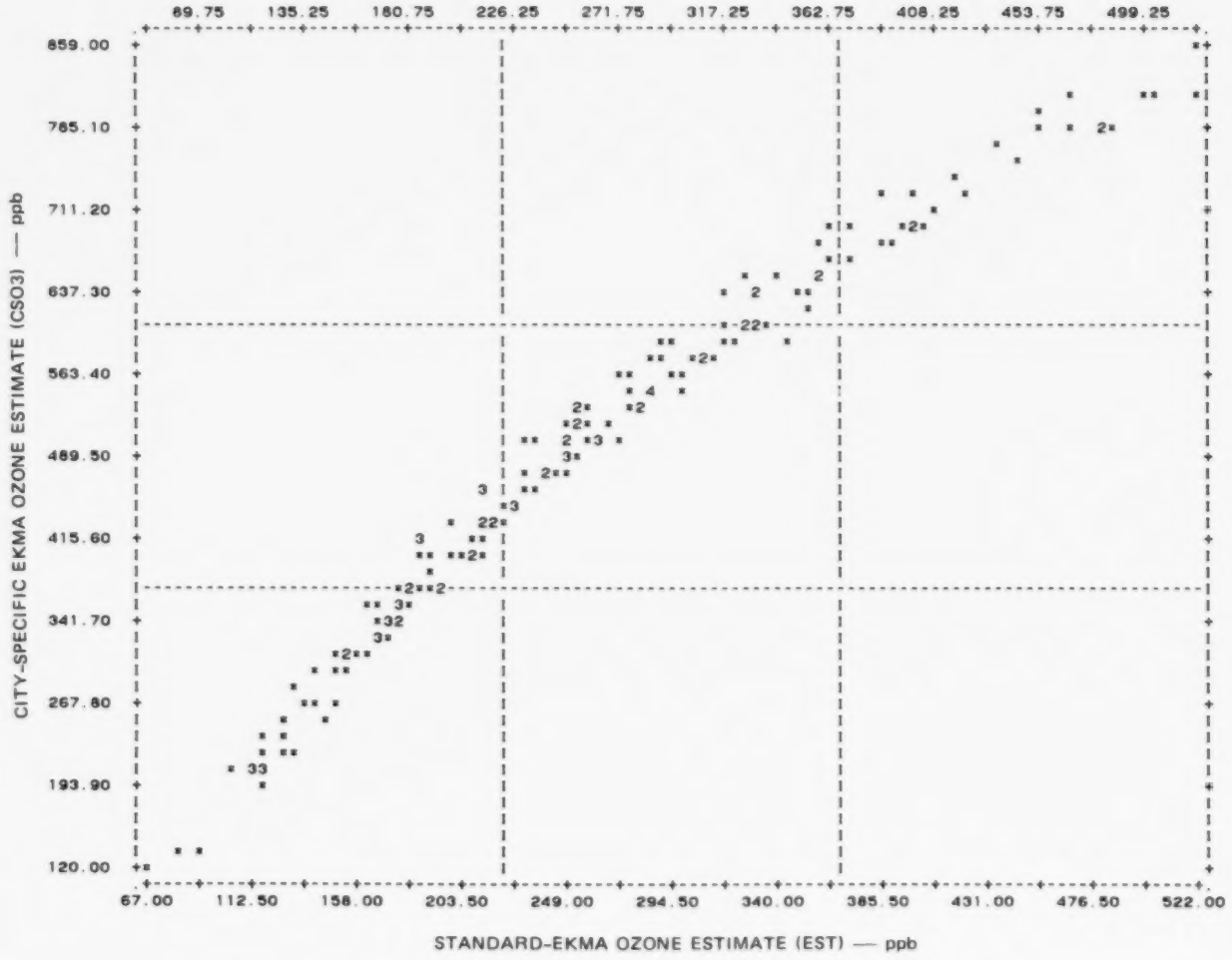


Figure 37. Scatterplot of city-specific EKMA ozone estimates and standard-EKMA ozone estimates for Los Angeles.

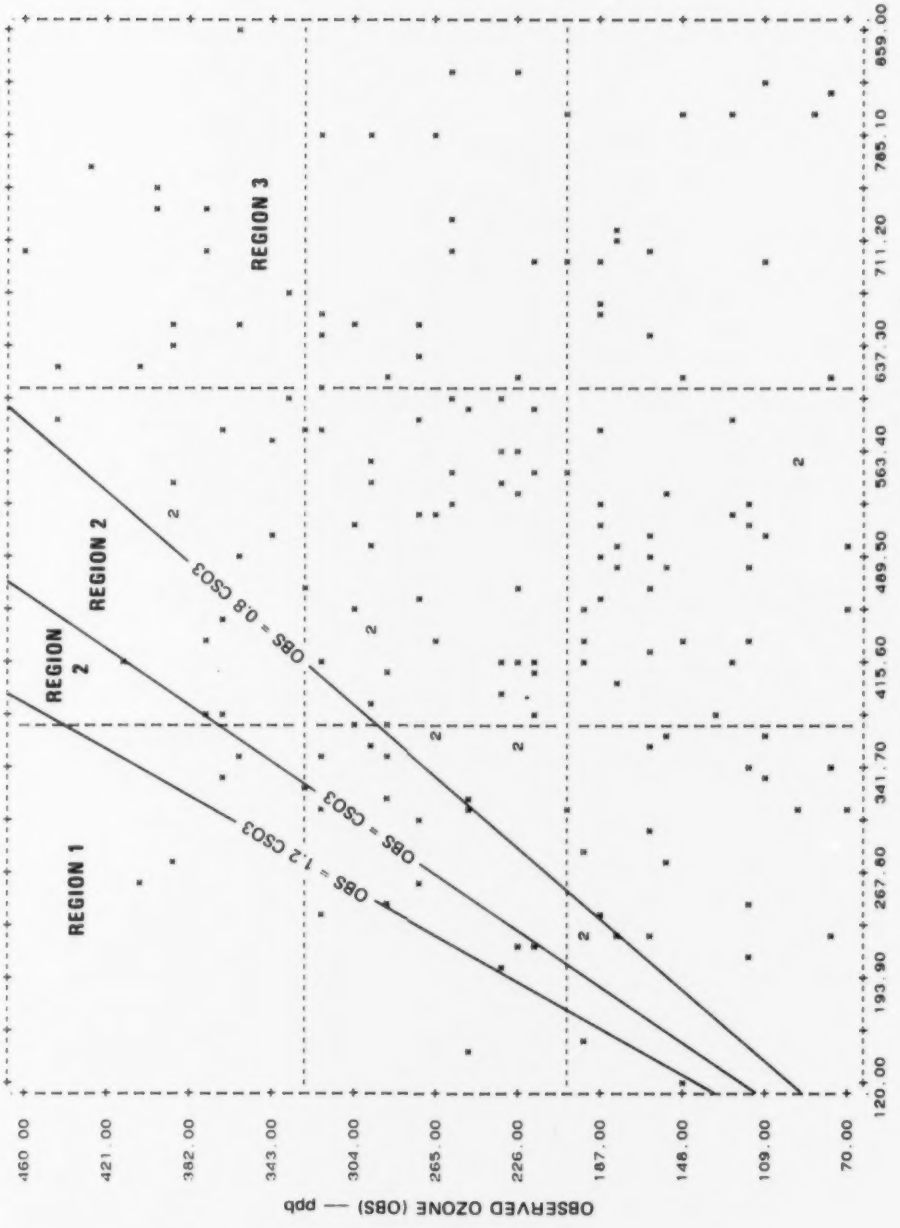


Figure 38. Scatterplot of observed ozone and city-specific ozone estimate for Los Angeles.

All the coefficients are statistically significant ($p < 0.001$), but the constant is not significant at the 0.05 level.

Equation 11 differs from Eq. 10 in the presence of the variable NO_x in the former instead of the NMOC that appeared in the latter. Nevertheless, in both equations 1/NMOC explained approximately the same amount of variance, 37 percent in Eq. 11 and 40 percent in Eq. 10. The remaining three variables in Eq. 11, $1/NO_x$, NO_x and T, together account for about 10 percent of the variance.

Figure 39 displays the plot of accuracy probability corresponding to Eq. 11 with $T = 76.1^{o}F$, which is its mean value. The probability is plotted as a function of the variable $Z = -54.999/NO_x + 286.822/NMOC - 0.001687(NO_x)$, which is part of the right side of Eq. 11. The figure shows that the probability of underprediction is low for most values of Z. The probability of an accurate prediction is moderately high, with $P(0.8 \leq R \leq 1.2)$ reaching a maximum value of approximately 0.64 in the neighborhood of $Z = 0.22$. The tendency to overpredict is apparent in the fact that $P(R < 0.8) > 0.5$ for $Z < 0$. The highest probabilities of an accurate prediction range from 0.6 to 0.64 and occur for values of Z in the interval $0.1 < Z < 0.3$.

Constant-Z contours are displayed on the NMOC-NO_x plane in Figure 40. The shaded area of Figure 40 corresponds to the interval $0.1 < Z < 0.3$ in Figure 39 in which $P(0.8 \leq R \leq 1.2) > 0.60$. Thus, (NMOC,$NO_x$) combinations in the shaded area have the highest probability of yielding accurate estimates. The small area to the left of the shaded slice is the region where underprediction becomes more probable. The part of the evaluation region to the right of the shaded area is associated with an increasing probability of overprediction. The tendency toward overprediction is thus made apparent, because most of the evaluation region is to the right of the shaded area.

DISCUSSION

It was surprising that the standard-EKMA yielded so many underestimates, indicating that the worst-case conditions supposedly embodied in the standard EKMA do not in fact define a worst case. The city-specific EKMA, by contrast, yielded a large majority of overestimates. This suggests that the city-specific EKMA is the operational mode of choice for the purpose of obtaining an upper bound for ozone, although the magnitude of the overprediction can be very large.

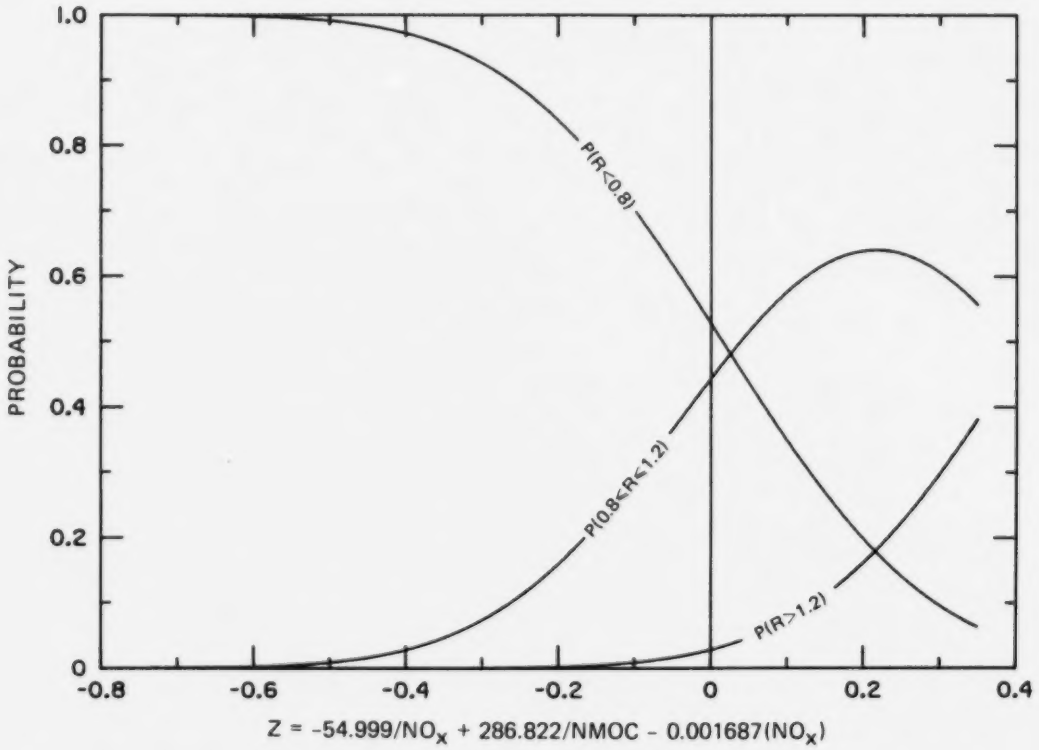


Figure 39. Accuracy probability plot for city-specific EKMA ozone estimates for Los Angeles. Mean value is assumed for maximum daily temperature.

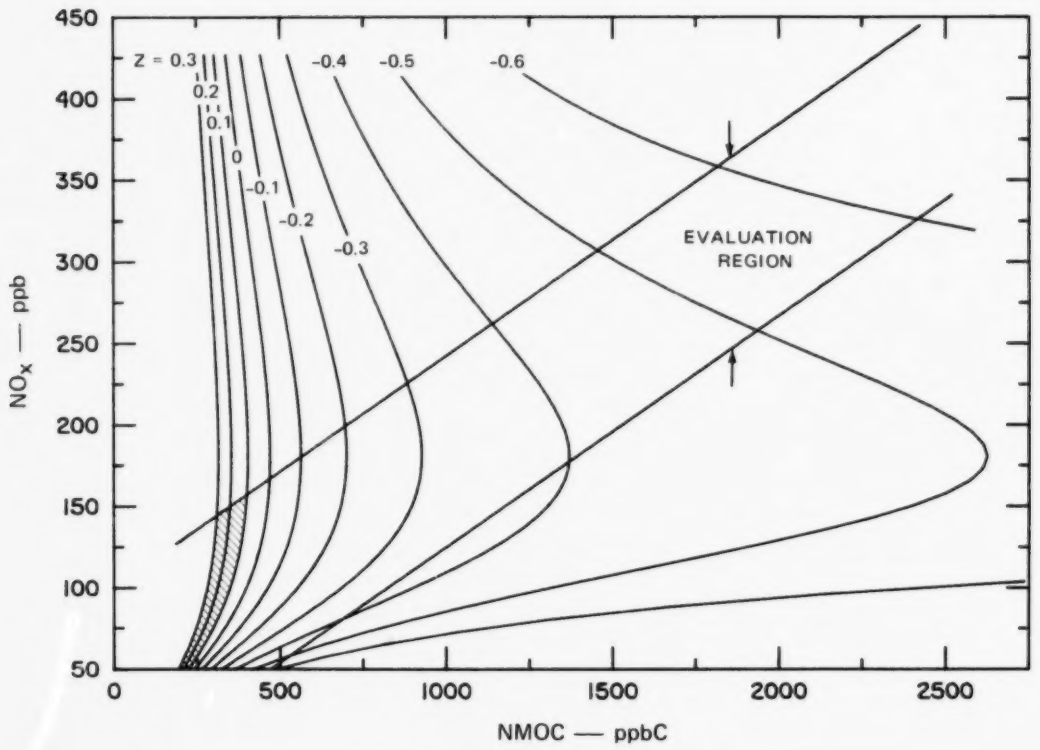


Figure 40. Plot of constant-Z curves on NMOC-NO_x plane for city-specific EKMA ozone estimates for Los Angeles. Shaded area denotes region where accurate estimates are most probable, assuming mean maximum daily temperature.

SECTION 7

EVALUATION USING DATA FOR THE TULSA AREA

DATA REVIEW AND ANALYSIS

Data collected during a field study conducted in the Tulsa, Oklahoma, area from July from September 1977 (Eaton et al., 1979) were used in the EKMA evaluation. Ten monitoring locations were used to observe the production and transport of ozone and ozone precursors, particularly when the winds were from the south (Figure 41).

In comparison with the other data bases used in the EKMA evaluation, peak ozone concentrations in the Tulsa area were generally lower and occurred later in the day. However, the peak ozone concentrations were significantly correlated with the observed daily maximum temperature, although higher temperatures are apparently needed to produce ozone concentrations exceeding the NAAQS ozone standard of 120 ppb. The 0600-0900 areawide-average source-region NMOC and NO_x concentrations are consistently lower in Tulsa than in the other test cities. Peak ozone concentrations were also significantly correlated with the morning-minimum to afternoon-maximum temperature difference.

NO_x and NMOC were measured at two monitors located within the Tulsa city limits. These two sites, the Tulsa Post Office and the Tulsa Health Department, were used to calculate the 0600-0900 average source-region NMOC and NO_x concentrations.

DEFINITION OF EVALUATION DATA SET

The criteria used to select the evaluation data set for Tulsa were:

- Maximum daily temperature at least 30.6°C
- Difference between daily maximum and morning minimum temperature at least 10.6°C.

Many days were excluded from the data set due to missing NMOC or NO_x data. One day with a maximum ozone level of 115 ppb was excluded because the maximum occurred before noon.

Thirteen days were selected fo the evaluation data set for Tulsa, which is shown in Table 19. A complete statistical evaluation of the EKMA, similar to that performed for the other four test cities, could not be performed for

SOURCE: Eaton et al. (1979)

Figure 41. Map of Tulsa, Oklahoma, and vicinity.

TABLE 19. EVALUATION DATA SET FOR TULSA

Date (1977)	Precursors		Temperature (°C)		Observed Maximum Ozone			EKMA Ozone Estimate (ppb)	
	NMOC (ppbC)	NO_x (ppb)	Daily Maximum	Morning Minimum	Site*	Time (CDT)	O_3 (ppb)	Standard	City-Specific
25 Jul	460	42	41.7	26.7	7	1900	123	147	101
29 Jul	440	53	34.4	23.3	1	1800	127	155	107
30 Jul	270	84	35.6	21.1	6	1300	128	108	85
2 Aug	960	65	35.0	19.4	1	1900	117	215	136
3 Aug	490	50	35.0	22.8	6	1500	151	159	108
4 Aug	500	59	35.0	23.9	7	1400	111	168	115
6 Aug	210	79	37.2	25.0	6	1600	79	74	70
7 Aug	220	107	38.2	25.0	6	1500	94	48	59
8 Aug	390	101	38.9	26.1	6	1500	84	156	109
9 Aug	410	76	38.3	25.0	6	1800	74	177	112
11 Aug	370	29	31.1	20.6	1	1600	95	120	85
14 Aug	170	10	33.9	21.7	8	1400	95	65	53
15 Aug	570	56	34.4	21.7	7	1300	106	174	117

*Site numbers are keyed to Figure 41.

Tulsa because of the small size of the data set. Consequently, we present an abbreviated analysis in the sections that follow.

EVALUATION OF STANDARD EKMA

Comparison of Observed and Estimated Ozone

Figure 42 shows a scatterplot of OBS and EST. Because the number of points is so small, the plot yields no statistically meaningful pattern. However, each of the three regions contains some points, and the respective number of points in Regions 2 and 3 differs only by one. Moreover, there is no statistically significant correlation between OBS and EST. The paucity of data prevents us from inferring any trends from Figure 42.

Relationship Between the Variables

A scatterplot of 0600–0900 NMOC and NO_x is shown in Figure 43. In contrast to the other four test cities, the NMOC and NO_x in Tulsa are not correlated. One possible implication of this lack of correlation is that the source region, if it exists, has not been properly characterized by the two monitoring stations used to compute average NMOC and NO_x. Another possible explanation is that stationary sources may be affecting one or both putative "source-region" monitors, which would weaken the correlation between NMOC and NO_x. In fact, Figure 43 shows three clusters of points, one comprising five points in the upper left hand corners of the graph, one seven points that are aligned along the figure's main diagonal, and one a single point located in the middle of the right-most margin. The seven-point cluster suggests a linear relationship between NMOC and NO_x, but the dichotomy between the five-point and seven-point clusters could be indicative of the presence of multiple uncorrelated sources affecting the precursor levels. It is thus apparent that the two monitors are inadequate to characterize the Tulsa source region, if indeed it exists.

No significant correlation was found between OBS and NMOC and NO_x. However, EST is strongly correlated with NMOC ($r = 0.87$, $p < 0.001$) and with 1/NMOC ($r = -0.94$, $p < 0.001$). But EST is not correlated with NO_x or $1/NO_x$, which is another indication of the possible influence of stationary sources on the 0600–0900 NO_x.

EVALUATION OF CITY-SPECIFIC EKMA

City-Specific Ozone Estimates

Table 20 shows the input parameters for the city-specific calculations. The trajectories calculated for estimating post-0800 emissions all left the source area after 0900; hence, there are no post-0900 emissions. The absence of post-0900 emissions is a consequence of the small size of the urban area and of the fact that Tulsa is surrounded by open lands.

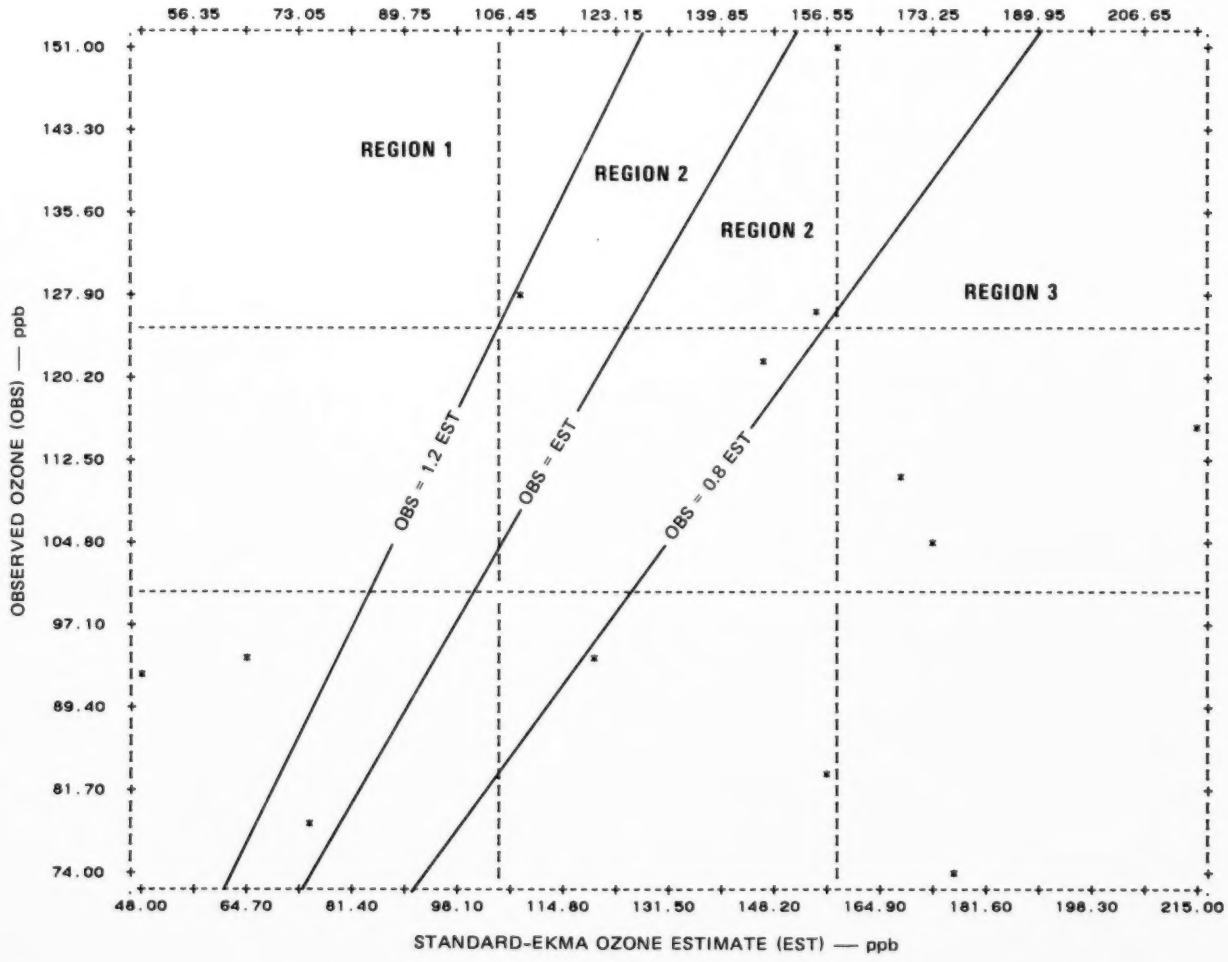


Figure 42. Scatterplot of observed ozone and standard-EKMA ozone estimate for Tulsa.

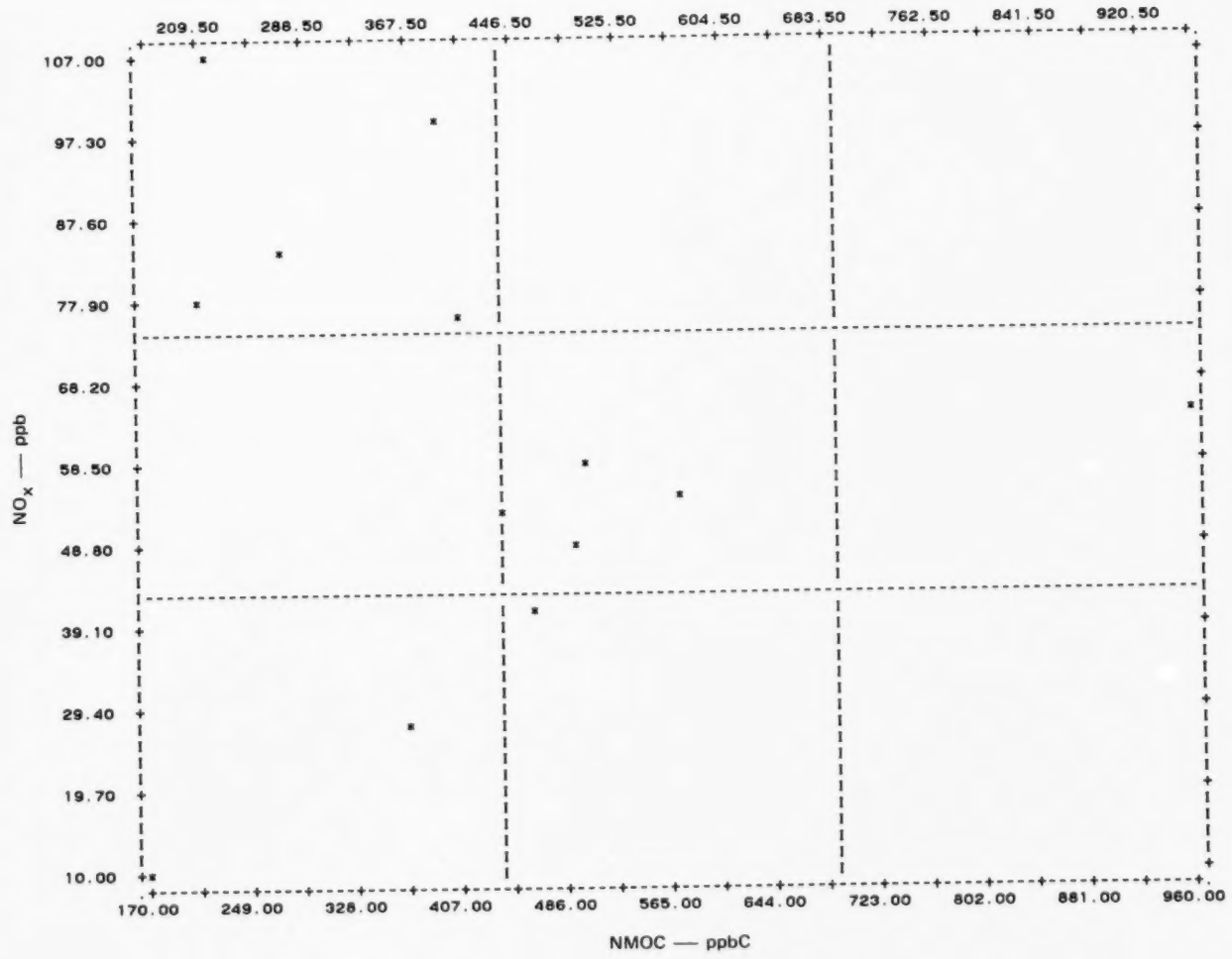


Figure 43. Scatterplot of 0600–0900 NMOC and NO_x for Tulsa.

TABLE 20. SUMMARY OF INPUT PARAMETERS FOR OBTAINING CITY-SPECIFIC EKMA OZONE ESTIMATES FOR TULSA

Parameter	Value
Date	3 August 1977
Location	Latitude 36.92°N, Longitude 95.92°W
Inversion height data	
Initial height (m)	400
Final height (m)	1800
Starting time of rise	0700
Ending time of rise	1400
Post-0800 emissions*	
NMOC	0.50
NO_x	0.50
Reactivity, NO_2/NO_x	0.52
Background ozone concentration (ppm)	0.04

*Ending hour 0900.

As in the other four test cities, the C-S ozone estimates are strongly correlated with the standard-EKMA predictions ($r = 0.99$, $p < 0.001$). The C-S estimates are also correlated with NMOC ($r = 0.88$, $p < 0.001$), with 1/NMOC ($r = -0.95$, $p < 0.001$), but not with NO_x or $1/NO_x$.

Comparison of Observed Ozone and City-Specific Ozone Estimates

Figure 44 shows the scatterplot of OBS and CS03. Comparison with Figure 42 shows that the C-S estimates display an increased frequency of underprediction with a corresponding decrease in overprediction. The frequency of accurate predictions remains about the same as for the standard-EKMA case.

DISCUSSION

Although the small sample size precludes drawing general conclusions about EKMA performance in the Tulsa area, a few cautionary remarks can nonetheless be made. In using the EKMA, it is important to have a reasonably well-defined source region. The Tulsa application revealed that the source region was either inadequately defined or it does not exist. In either case, the use of EKMA in the Tulsa context would be inappropriate. Thus, any further attempts to apply the EKMA to Tulsa should be preceded by an investigation of the source distribution, with the aim of determining whether the source-region concept that underlies the EKMA fits Tulsa's conditions.

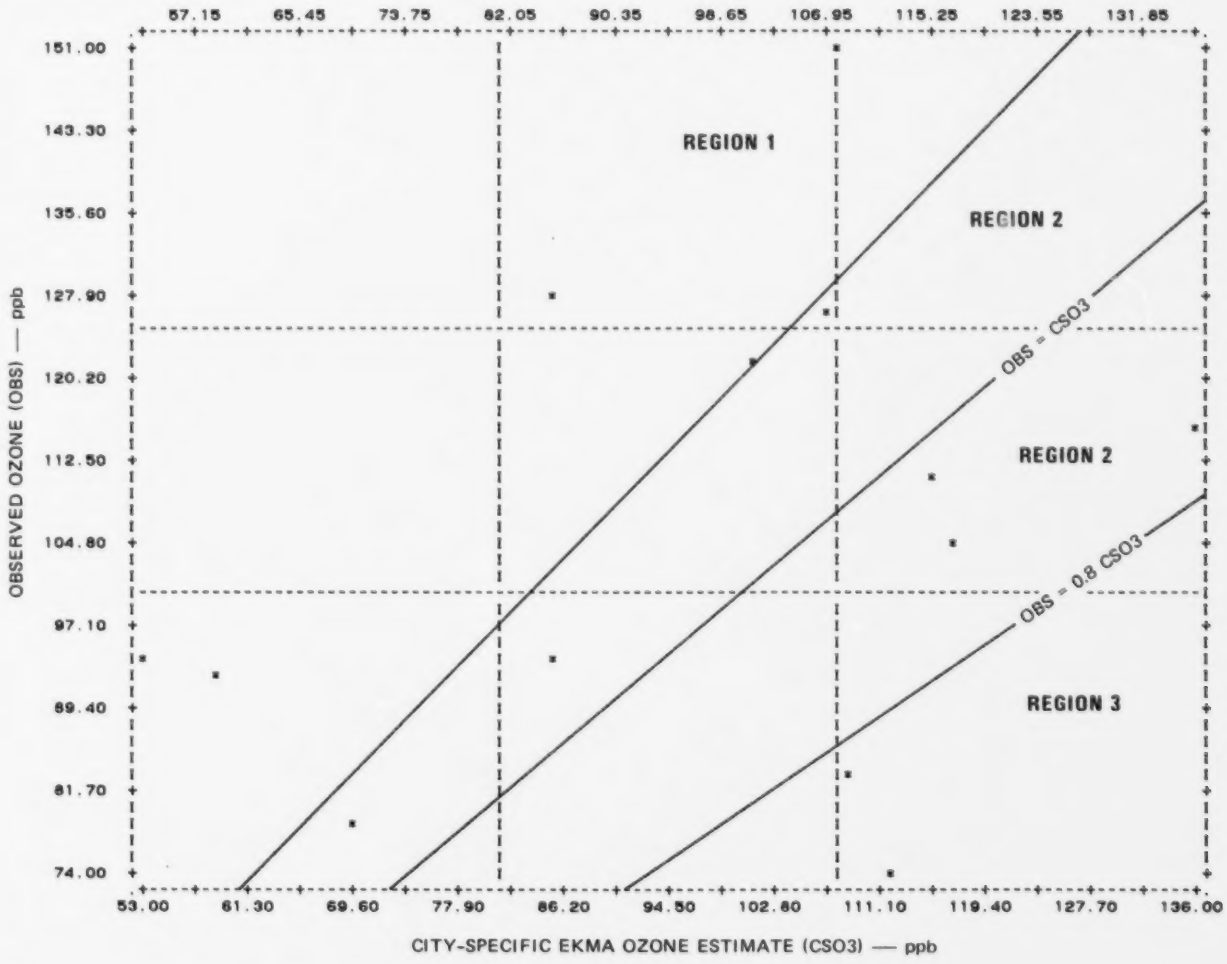


Figure 44. Scatterplot of observed ozone and city-specific EKMA ozone estimate for Tulsa.

SECTION 8

CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS

DISCUSSION OF RESULTS

This study has:

- Demonstrated the feasibility of using the EKMA to estimate upper bounds for daily maximum ozone, given the concentration of NMOC and NO_x.
- Shown that the accuracy of the EKMA ozone estimates is a function of the position of the estimate on the O_3 isopleth diagram, the position being determined by the (NMOC, NO_x) coordinates.
- Developed a general method for evaluating EKMA performance as a predictor of maximum ozone.

The results indicate that it is possible to use the EKMA, in either the standard or city-specific form, to estimate an upper bound for daily maximum ozone, given the concentration of NMOC and NO_x. However, for a certain range of values of NMOC and NO_x the upper bound can be very loose, and thus of limited usefulness. The EKMA also can produce ozone estimates that are accurate to within ±20 percent of the observations. Estimates of this accuracy can be obtained only for a restricted range of values of NMOC and NO_x. In some cases, e.g. for St. Louis, the range of values of NMOC and NO_x that determines the accuracy of the estimates is dependent in turn on such other variables as background ozone level and maximum daily temperature.

Three levels of accuracy were defined based on the ratio R = observed/estimated: $R > 1.2$, $0.8 \leq R \leq 1.2$, and $R < 0.8$. The interval $R > 1.2$ defines cases of underestimation; $R < 0.8$ defines cases of overprediction. The closed interval $0.8 \leq R \leq 1.2$ defines the most accurate estimates. The accuracy of the ozone estimates depended not only on the level of pollutant and meteorological variables, but also on whether the standard or city-specific EKMA was used. Moreover, the frequency of occurence of ozone estimates that fall in each of the three accuracy intervals was different for all the data sets studied. Nevertheless, a general pattern emerged that relates low, medium, and high values of NMOC and NO_x to the three accuracy intervals: Low values tended to yield $R > 1.2$, medium values, $0.8 \leq R \leq 1.2$, and high values, $R < 0.8$. The precise values of NMOC and NO_x that mark the boundaries of the three accuracy intervals differed among the individual data sets. St. Louis and Houston exhibited a general trend toward a high frequency of overestimation, and a low frequency of underprediction, for the standard EKMA. Although for St. Louis and Houston the standard EKMA yielded estimates in the interval $0.8 \leq R \leq 1.2$,

the frequency was low, and the range of NMOC and NO_x values that produced such estimates with high probability was narrow.

The situation was different for the city-specific EKMA, which tended to have a higher frequency of estimates in the interval $0.8 \leq R \leq 1.2$. The city-specific EKMA also tended to underpredict more frequently, and to overpredict less frequently, than the standard EKMA. Thus, although the city-specific EKMA offers a higher probability of producing an accurate estimate (for which $0.8 \leq R \leq 1.2$), the standard EKMA has a higher probability of yielding an upper bound estimate (for which $R \leq 1.2$). This has implications for EKMA applications that are discussed in the next section.

The situation described above for St. Louis and Houston was reversed for Philadelphia and Los Angeles. For the latter two test cities the standard EKMA had a higher frequency of underprediction and a lower frequency of overprediction than the city-specific EKMA. Consequently, the city-specific EKMA is more useful for obtaining upper bounds on maximum ozone potential for Philadelphia and Los Angeles.

A major thrust of the study was to quantify the probable error associated with the estimates. Equations were derived for calculating the probability that the observed/estimated ratio for a given estimate was in one of the three accuracy intervals, namely $R > 1.2$, $0.8 \leq R \leq 1.2$, and $R < 0.8$. Depending on the data set studied, the probability was a function of NMOC and/or NO_x and other variables. Accuracy regions were identified in the NMOC-NO_x plane wherein an estimate can be assigned a probability of having attained a prescribed accuracy level. Although our computations were aimed at the three specific accuracy levels previously mentioned, the equations derived are general and can be applied to any arbitrarily defined accuracy levels.

The study's use of several distinct data bases lends generality to the patterns and trends discussed above. Nevertheless, the scope of the numerical results is limited, in the sense that they strictly apply only to the particular data sets examined. Unless these data sets can be considered to be representative of conditions in their respective geographical areas, the application of the equations and graphs derived in this study to other geographical regions, or to the same regions in other years, must be regarded with caution. Such restrictions are, of course, typical of most data-analytic studies.

We regard the evaluation methodology developed in the study as an important contribution. The methods are general, and the numerical results reported are real-life examples of what can be accomplished with these techniques. In the future, the same methodology can be applied to other data bases to extend and generalize further the results reported here, as recommended below.

IMPLICATIONS FOR EKMA APPLICATIONS

Using the EKMA to estimate maximum ozone from its NMOC and NO_x precursors implies that the EKMA could be used to assess the effect on ozone of actions

that modify the concentration of NMOC and NO_x, e.g. emission control strategies. A corollary is that the EKMA could also be used in the design of control strategies by performing a sequence of analyses of the effect on ozone of a variety of postulated control measures. Below we examine the possibilities and pitfalls of using the EKMA in this fashion in light of the results of this study. The discussion will be cast in terms of emission control strategies, because we consider this to be the most common application of the EKMA. However, it should be understood that any action that modifies precursor levels is implicitly treated. The discussion assumes that the reader is familiar with the principles and assumptions of the EKMA as presently formulated.

The problem at issue is to evaluate the effectiveness of a control strategy aimed at reducing ozone by curtailing emissions of NMOC or NO_x, or both. In this context, the EKMA could be used to estimate the maximum ozone associated with the control strategy in an attempt to answer the following questions:

(1) What is the probability that the maximum ozone will exceed 120 ppb?

(2) If it appears that 120 ppb will be exceeded, is the estimated ozone close to or much greater than 120 ppb?

The general procedure is depicted in the flow chart in Figure 45. The EKMA could be used to answer the first question provided that the ozone estimate is less than or equal to 120 ppb, and there is a high probability that the estimate is an upper bound for the actual concentration. This corresponds to taking the right-hand branch of Figure 45. In this case, the answer to Question 1 above is that the probability is very low that the control strategy will produce ozone levels that exceed 120 ppb; the precise value of the probability would be obtained from a probability graph such as Figure 7. Note that the accuracy of the ozone estimate is not important in this situation because an upper bound estimate is all that is needed. Hence, the standard EKMA may be satisfactory for this application.

The EKMA is not as helpful in answering Question 1 if the estimate is under 120 ppb but there is a high probability of underprediction. As indicated in the right-most branch of Figure 45 in this case, the analysis must be refined to establish the accuracy of the estimate and the probability associated with that accuracy. The methods developed in this study allow a user to perform such a refined analysis. Because accuracy is important in these circumstances, it may be necessary to resort to the city-specific EKMA to obtain the ozone estimate. However, this must be tempered by the knowledge that the city-specific EKMA has a tendency to underpredict.

Answering Question 2 requires analyzing the accuracy of the estimate and its associated probability. As depicted in the left branch of Figure 45, three cases are to be considered when the estimate is over 120 ppb:

(1) There is a high probability of overprediction

(2) There is a high probability of underprediction

(3) There is a relatively high probability that the estimate is accurate, for example, to within ±20 percent.

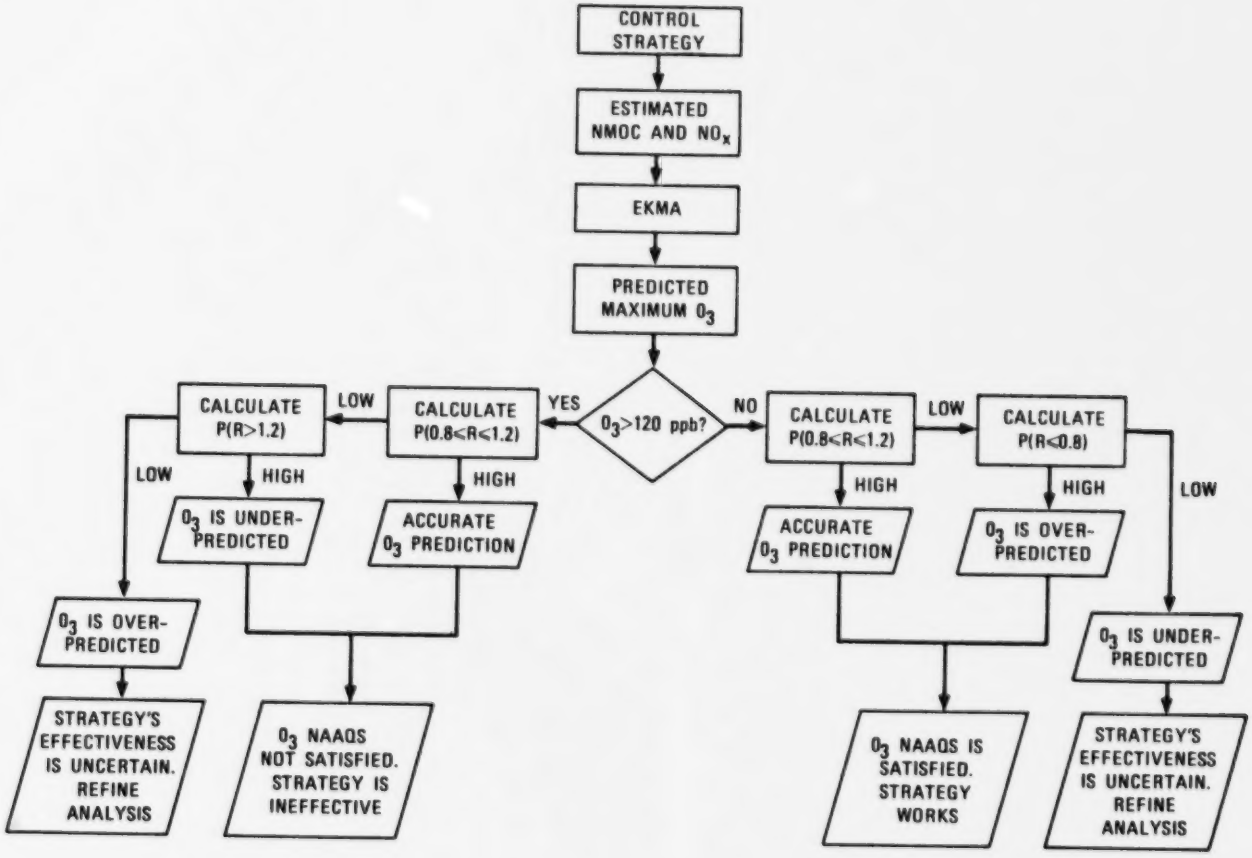


Figure 45. Flowchart for the application of the EKMA as a screening tool for analyzing the impact of a control strategy.

The EKMA is least helpful in the first case (the left-most branch of Figure 45) because although the control strategy may actually reduce ozone below 120 ppb, the overprediction masks the effect. This case thus requires a very thorough analysis of the accuracy of the estimate. In the second case, it is clear that the control strategy does not work, and the margin of ineffectiveness should be assessed by analyzing the accuracy of the estimate as a means of guiding the reformulation of the control strategy. The third case is where the EKMA is most useful, because the estimate is relatively accurate. Such an estimate can also be used to guide the design of a new control strategy.

Used in the fashion described above, the EKMA can be considered to be a screening tool, albeit coarse at times, that allows one to analyze the potential impact of a control strategy. Coupled with the analysis of the accuracy of the estimates, the EKMA could also be used to help formulate a control strategy by sequentially screening a series of control strategies. In general, we recommend applying the standard EKMA first, then going to the city-specific mode if the results obtained with the standard mode warrant it.

To apply the EKMA in the manner described, one must have a means of calculating the accuracy, and its associated probability, of the ozone estimate. This, of course, presumes that EKMA performance has been already evaluated following the methods described in this report. Hence, one would have available probability curves similar to Figure 7 and equations such as Eq. 2 for analyzing the accuracy of the estimates. Because such curves and equations would be based on a particular set of existing data, one is faced with the perennial problem of deciding whether it is appropriate to use them for predictive purposes; in effect, to extrapolate using equations and graphs based on historical data. This judgment is not unique to this EKMA application; all models are tested and verified using historical data and are then applied predictively. The prudent course is not to stray too far from the known conditions.

RECOMMENDATIONS

We propose three extensions of the current research, first that further evaluations of the EKMA be conducted using the methods developed in this study. Specifically, evaluations should be performed using data for other years in St. Louis, Houston, Philadelphia, and Los Angeles. This suggestion is motivated by the need to establish whether the pattern of EKMA performance associated with a single data set for each of these areas is general in nature. Only these four urban areas are recommended, because the Tulsa data base that is normally available is not sufficiently detailed for such analysis. Adding some new areas to the extended evaluation should also be considered, provided that at least two years of data are used for each area.

Second, we recommend investigating the application of the results of this study to the analysis and design of ozone control strategies. A procedure can be developed that uses EKMA ozone estimates and associated accuracy probabilities to analyze the effectiveness of a control strategy with respect to its effect on ozone levels. The analysis can be converted into design by sequentially examining the impact of a series of control strategies. The results of

the studies described in this report lend themselves to the use of Monte Carlo simulation techniques in conjunction with the EKMA to predict the distribution of ozone maxima. This approach uses the entire joint distribution of NMOC and NO_x as the model input and produces the distribution of ozone maxima as output. Thus, EKMA applications would not be limited to using only the median 0600-0900 NMOC/NO_x ratio and the single observed ozone design value as the sole EKMA inputs.

The third recommendation is prompted by the finding that the standard and city-specific EKMA ozone estimates are highly correlated. The linear relationship between the two sets of estimates indicates that the city-specific estimates are scaled versions of the standard estimates and the standard scaled versions of the city-specific. The presence of a nonzero constant term in the linear equation indicates that the two sets of ozone isopleths are shifted with respect to one another. The well-defined linear transformation between the two sets of estimates suggests that the transformation can perhaps be generalized and parameterized so that the city-specific estimates can be obtained directly from the standard estimates. This would eliminate the necessity of running the EKMA computer code in the city-specific mode, which would simplify considerably the use of EKMA: equations already exist that fit the standard EKMA ozone isopleth diagram (see the appendix of this report). Hence, a pocket calculator could be used to compute the estimated ozone for both standard and city-specific EKMA. Thus, it is recommended that EPA sponsor an effort to derive a general transformation between standard and city-specific EKMA ozone estimates.

CONCLUDING REMARKS

The model evaluation methodology developed in this study has several valuable features, namely:

- It is independent of the model evaluated. Hence it can be used with other models besides EKMA.
- The accuracy of the model predictions can be calculated from the model inputs.
- Accuracy bounds can be expressed in terms of probabilities.

These attributes make the methodology useful for application to a variety of model evaluation studies. The methodology is especially easy to apply in connection with trajectory models such as EKMA or ELSTAR. Such models have a relatively small set of inputs, and this facilitates the derivation of the regression equation that describes the predictive behavior of the model. Hence, we urge that the methodology be considered for use in other model evaluation studies. For example, it could be applied to Gipson and Meyer's (1981) recent study.

APPENDIX

EQUATIONS FOR STANDARD-EKMA OZONE ESTIMATES

The standard-EKMA ozone estimates were obtained using a set of equations fitted to the ozone isopleth diagram by researchers at the University of North Carolina (Holton, 1980; Holton and Jeffries, 1979). An independent test of the equations was performed at SRI for this study that detected a slight bias in the equations. The bias was corrected by means of adjustment equations that we derived using regression analysis. The ozone isopleth equations and the correction expressions are defined below.

Following the notation of Holton and Jeffries (1979), let

$$\theta = \tan^{-1} (1/5.49)$$

$$L = (\text{NMOC}) \cos \theta + (NO_x) \sin \theta$$

$$D = (\text{NMOC}) \sin \theta - (NO_x) \cos \theta$$

For $D \geq 0$, the uncorrected ozone is given by

$$O'_3 = 0.282\ L^{0.683}\ [1 - (5.49\ D/L)^{1.97}\ ]^{0.518} \quad .$$

The corrected ozone is given by

$$O_3 = 0.00338337 + 1.00075\ (O'_3) \quad .$$

For $D < 0$, the uncorrected ozone is given by

$$O''_3 = 0.282\ L^{0.683}\ \exp\ [-9.11\ (|D|/L)^{1.68}\ ] \quad .$$

The corrected value is obtained from

$$O_3 = 0.00912572 + 1.08471\ (O''_3) \quad .$$

REFERENCES

Cardwell, G., 1980: "Regional Air Pollution Study-Gas Chromatography Laboratory Operation," EPA-600/4-80-006 (January).

Dodge, M.C., 1977: "Effect of Selected Parameters on Predictions of a Photochemical Model," EPA-600/3-77-048 (June).

Eaton, W.C., C.E. Decker, J.B. Tommerdahl, and F.E. Dimmock, 1979: "Study of the Nature of Ozone, Oxides of Nitrogen, and Nonmethane Hydrocarbons in Tulsa, Oklahoma," EPA-450/4-79-008a.

Eaton, W.C., and F.E. Dimmock, 1979: "Study of the Nature of Ozone, Oxides of Nitrogen, and Nonmethane Hydrocarbons in Tulsa, Oklahoma, Volume II: Data Tabulation," EPA-450/4-79-008b (April).

Gipson, G.L., and E.L. Meyer, 1981: "Simplified Trajectory Analysis Approach for Evaluating OZIPP/EKMA," presented at EKMA Workshop, Research Triangle Park, North Carolina (15 December).

Holton, G.A., 1980: Private communication, University of North Carolina, Chapel Hill (16 September).

Holton, G.A., and H.E. Jeffries, 1979: "Mathematical Analysis of Ozone Isopleths," Proceedings of Specialty Conference on Ozone/Oxidants Interactions with the Total Environment, Houston, Texas (14-17 October).

Ludwig, F.L., and J.R. Martinez, 1979: "Aerometric Data Analysis for the Houston Area Oxidant Study--Volume II," SRI International, Menlo Park, California, Project 7106 (July).

Ludwig, F.L., J.R. Martinez, and K.C. Nitz, 1979: "Data Summaries for the Houston Area Oxidant Study," Houston Area Oxidant Study Contact DA-1, SRI International, Menlo Park, California.

Martinez, J.R., 1982: "Characterization of Data Quality for the 1978 Houston Oxidant Modeling Study," Final Report, SRI Project 7938, SRI International, Menlo Park, California (March).

Maxwell, C., 1981: "Empirical Relationships of Solar Radiation and Meteorological Data to Ozone Concentrations," presented at the Fifth Symposium on Turbulence, Diffusion, and Air Pollution, American Meteorological Society, Boston, Massachusetts.

Nitz, K.C., and J.R. Martinez, 1980: "1978 Houston Oxidant Modeling Study; Volume II: Data Base Guide," EPA Contract 68-02-2984, SRI International, Menlo Park, California.

Nitz, K.C., and J.R. Martinez, 1982: "1979 Houston Oxidant Modeling Study, Volume II: Guide to the Data Base," Final Report, SRI Project 7938, SRI International, Menlo Park, California (March).

Strothman, J.A., and F.A. Schiermeir, 1979: "Documentation of the Regional Air Pollution Study (RAPS) and Related Investigations in the St. Louis Air Quality Control Region," EPA-600/4-79-076 (December).

Trijonis, J., and D. Hunsaker, 1978: "Verification of the Isopleth Method for Relating Photochemical Oxidant to Precursors," EPA-600/3-78-019 (February).

U.S. Environmental Protection Agency, 1977: "Uses, Limitations, and Technical Basis of Procedures for Quantifying Relationships Between Photochemical Oxidants and Precursors," EPA-450/2-77-021a (November).

U.S. Environmental Protection Agency, 1978: "Procedures for Quantifying Relationships Between Photochemical Oxidants and Precursors: Supporting Documentation," EPA-450/2-77-021b (February).

U.S. Environmental Protection Agency, 1980: "Guideline for Use of City-Specific EKMA in Preparing Ozone SIPs," EPA-450/4-80-027 (October).

Westberg, H., and P. Sweeny, 1980: "Philadelphia Oxidant Data Enhancement Study: Hydrocarbon Analysis," Washington State University, Research Report 80-13-36 (25 August).

Whitten, G.Z., and H. Hogo, 1978: "User's Manual for Kinetics Model and Ozone Isopleth Plotting Package," EPA-600/8-78-014a (July).